

TESIS DOCTORAL

**EVALUACIÓN DINÁMICA
DE PROCESOS IMPLICADOS
EN LA LECTURA**

Juan José Navarro Hidalgo

Director de la Investigación: Prof. Dr. *Joaquín Mora Roche*

Departamento de Psicología Evolutiva y de la Educación
Universidad de Sevilla

MEMORIA DE INVESTIGACIÓN PARA
LA OBTENCIÓN DEL GRADO DE DOCTOR

**EVALUACIÓN DINÁMICA
DE PROCESOS IMPLICADOS
EN LA LECTURA**

Juan José Navarro Hidalgo

Director de la Investigación: Prof. Dr. *Joaquín Mora Roche*

Departamento de Psicología Evolutiva y de la Educación
Universidad de Sevilla

Agradecimientos

El presente trabajo de investigación ha contado con la valiosa colaboración de un buen número de docentes que desempeñan su labor en los Centros educativos: CEIP Andalucía (Sevilla), CEIP Antonio Machado (Morón de la Frontera), CEIP Lope de Rueda (Sevilla), CEIP Manuel Sánchez Alonso (Arahal), CEIP El Ruedo (Arahal), Colegio Concertado Santa Isabel (Sevilla), CEIP Vicente Neria (Sevilla), IES Los Azahares (La Rinconada), IES La Jarcia (Puerto Real), IES Doña Leonor de Guzmán (Alcalá de Guadaíra), IES Mariana Pineda (Montequinto-Dos Hermanas), IES Torre de los Guzmanes (La Algaba), IES Torreblanca (Sevilla); a todos ellos quiero expresar mi agradecimiento y mi consideración por el tiempo, el interés y el esfuerzo empleados en las tareas encomendadas para este estudio.

De forma muy especial, quiero agradecer aquí su inestimable colaboración en el trabajo de campo de la investigación, a los docentes y orientadores que conformaron el *Grupo de Trabajo* a través del cual se ha podido llevar a la práctica la fase experimental de nuestra propuesta de evaluación dinámica, así como a aquellos que han participado desde sus Centros facilitando la aplicación de cuantas pruebas han sido necesarias: *María José Alés, Eva María Cortés, Helena Lama, Ángela Molina, Inmaculada Moreno, Gonzalo Vidal, Sandra Cantillo, Maribel Gallardo, Antonio Liñán, Nuria Montilla, Mónica Rodríguez, Rafael Anaya, Ana Lara, Joaquín Lara y Eva Sánchez*. Junto a estos profesionales, quiero también agradecer su colaboración y su implicación en las tareas de investigación encomendadas a los docentes: *Cinta Álvarez, Macarena Belchi, Anabel López, Moisés Luque, Janet Portillo y Jesús Manuel Ramírez*.

Quiero expresar igualmente mi gratitud al CEP de Alcalá de Guadaíra, y especialmente a *José Antonio Jiménez*, que acogió y facilitó la creación de un *Grupo de Trabajo* con el objetivo de aplicar y analizar la puesta en práctica de nuestra propuesta.

Por último, mi más profundo agradecimiento al profesor Dr. Joaquín Mora. Los momentos de encuentro con él han sido claves para la gestación y el desarrollo de este trabajo, y siempre han ocasionado la generación de *conflictos cognitivos activadores* de nuevas ideas o de transformaciones que mejoraban sustancialmente nuestro trabajo.

EVALUACIÓN DINÁMICA DE PROCESOS IMPLICADOS EN LA LECTURA

ÍNDICE

Introducción /1

Marco Teórico

1. Naturaleza y evaluación de la lectura /7

1.1 La lectura como construcción socio-histórica y dinámica

Abordaje conceptual de la lectura: características básicas de los modelos explicativos (8)

Estrategias y procesos contemplados en la lectura (11)

1.2 La evaluación de la lectura y la comprensión de textos (17)

Introducción (17)

Enfoques evaluativos: El enfoque del test y la evaluación integrada en el proceso de enseñanza y aprendizaje (19)

La evaluación de la lectura y la comprensión desde el enfoque del test (20)

La evaluación de la lectura y la comprensión desde el enfoque integrador (24)

Técnicas y estrategias específicas de evaluación del proceso de comprensión (29)

La entrevista (29)

El registro del pensamiento en voz alta (30)

La observación (30)

La tutorización entre iguales (33)

Detección de incoherencias lingüísticas (33)

Medidas directas del proceso de comprensión (34)

Técnicas y estrategias específicas de evaluación del producto de la comprensión (34)

Prueba de recuerdo libre (34)

Recuerdo provocado (35)

Resumen (35)

Evaluación mediante preguntas sobre el texto (36)

Pruebas de cierre o completamiento (38)

Criterios de evaluación utilizados en la evaluación de la comprensión (39)

1.3 La metacognición y las dificultades de aprendizaje en la evaluación de la comprensión (43)

Metacognición y Dificultades en el aprendizaje (44)

Comprensión de textos y Metacognición (48)

Metaconocimiento y autorregulación del proceso de comprensión (49)

La planificación (50)

El control y la supervisión del proceso de lectura y comprensión (51)

La autoevaluación (53)

Procesos de transferencia y generalización (54)

La evaluación de la metacompreensión lectora. Algunos estudios sobre evaluación de procesos metacognitivos implicados en la comprensión (56)

2. Evaluación dinámica /59

- 2.1 Contextos de desarrollo potencial y conceptos clave en la evaluación dinámica (59)
 - Introducción (59)
 - Interacción mediada como concepto clave en la evaluación dinámica (60)
 - Los instrumentos psicológicos y la mediación semiótica (60)
 - La experiencia de aprendizaje mediado (63)
 - Las aptitudes como formas de experiencia en desarrollo (66)
- 2.2 Aproximación a las pruebas dinámicas y a los diseños metodológicos de referencia (67)
 - La necesidad de una evaluación dinámica (69)
- 2.3 Modelos y propuestas de evaluación dinámica (72)
 - Introducción (72)
 - La evaluación del Potencial de Aprendizaje de R. Feuerstein (73)
 - Bases teóricas y metodológicas (73)
 - Investigación y revisión crítica del modelo (77)
 - El dispositivo de evaluación del aprendizaje en preescolar (PLAD) de C.S. Lidz y C. Thomas (83)
 - Bases teóricas y metodológicas (83)
 - Investigación y revisión crítica del modelo (83)
 - La Batería de Modificabilidad Cognitiva (CMB) de D. Tzuriel (84)
 - Bases teóricas y metodológicas (84)
 - Investigación y revisión crítica del modelo (84)
 - El modelo de evaluación dinámica MindLadder, de M.R. Jensen (87)
 - Bases teóricas y metodológicas (87)
 - Investigación y revisión crítica del modelo (89)
 - El modelo de evaluación dinámica basada en el currículum (ACFS), de C.S. Lidz (93)
 - Bases teóricas y metodológicas (93)
 - Investigación y revisión crítica del modelo (96)
 - La evaluación en la Zona de Desarrollo Potencial y el método de las sugerencias graduadas de J. C. Campione y A. Brown (99)
 - Bases teóricas y metodológicas (99)
 - Investigación y revisión crítica del modelo (101)
 - La evaluación del potencial de aprendizaje a través del razonamiento inductivo (LIR) de Wilma C.M. Resing (104)
 - Bases teóricas y metodológicas (104)
 - Investigación y revisión crítica del modelo (107)
 - La Prueba del Potencial de Aprendizaje de M. Budoff (109)
 - Bases teóricas y metodológicas (109)
 - Investigación y revisión crítica del modelo (111)
 - La Evaluación del Potencial de Aprendizaje (EPA), de R. Fernández-Ballesteros (114)
 - Bases teóricas y metodológicas (114)
 - Investigación y revisión crítica del modelo (117)
 - Las pruebas de aprendizaje de J. Guthke (122)
 - Bases teóricas y metodológicas (122)
 - Investigación y revisión crítica del modelo (126)
 - La Prueba de Procesamiento cognitivo de H.L. Swanson (TPC-S) (129)
 - Bases teóricas y metodológicas (129)
 - Investigación y revisión crítica del modelo (133)
 - El modelo de Prueba de los Límites de J. Carlson y K.H. Wiedl (137)
 - Bases teóricas y metodológicas (137)
 - Investigación y revisión crítica del modelo (139)
 - La Batería de Evaluación Dinámica de Procesos Metacognitivos de D. Saldaña (145)
 - Bases teóricas y metodológicas (145)
 - Investigación y revisión crítica del modelo (148)

- Pruebas dinámicas para revelar el potencial de aprendizaje oculto. Los estudios de R. Sternberg y sus colaboradores (152)
 - Bases teóricas y metodológicas (152)
 - Investigación y revisión crítica del modelo (153)
- La Evaluación Dinámica de la competencia lingüística en una segunda lengua (154)
 - La Evaluación Dinámica de la Comprensión de Textos en una Lengua Extranjera, de A. Kozulin y E. Garb (154)
 - La Evaluación Dinámica de la aptitud para aprender otro idioma: la prueba CANAL-FT de Sternberg y Grigorenko (156)
- 2. 4 La evaluación dinámica de la lectura. Algunas experiencias que reabren el camino (159)
 - Una aplicación práctica con relación al dispositivo de evaluación presentado. La evaluación dinámica de metacimientos sobre la comprensión de textos (160)

3. La evaluación dinámica de procesos implicados en la lectura. Bases teóricas y metodológicas de nuestra propuesta y objetivos de la investigación /165

- 3.1 Bases teóricas y metodológicas de nuestra propuesta (165)
 - Bases teóricas y metodológicas generales de la evaluación dinámica de la lectura:
 - Cuestiones teórico-prácticas y psicométricas de base (165)
 - Aprendizaje y evaluación de la lectura: La relevancia del contexto social e histórico (173)
 - Evaluación de procesos: algunos problemas y propuestas de innovación en la evaluación de la comprensión (175)
 - La necesidad de una evaluación dinámica de los procesos de comprensión y del proceso de lectura y comprensión (177)
 - Una perspectiva metacognitiva. Nuestra propuesta de evaluación dinámica de procesos metacognitivos en la comprensión de textos (180)
 - La evaluación del contexto de aula (181)
 - Contextualizar los instrumentos de evaluación (185)
- 3.2 La evaluación dinámica de procesos implicados en la lectura. Objetivos de la presente investigación (187)

Método

4. Diseño de investigación e hipótesis planteadas /191

- 4.1 El diseño de investigación (191)
- 4.2 Hipótesis de investigación planteadas (198)
- 4.3 Colaboradores (204)
 - Formación de los colaboradores/aplicadores del dispositivo EDPL (207)

5. Participantes /209

- 5.1 Composición y distribución de la muestra (209)
- 5.2 Dificultades en la composición final de la muestra. Las características de los sujetos participantes y las decisiones adoptadas (211)
- 5.3 Análisis de la equivalencia entre los grupos (216)
 - Equivalencia con relación al sexo (217)
 - Equivalencia de edad y nivel escolar (220)
 - Equivalencia de nivel intelectual (222)
 - Equivalencia con relación al rendimiento en lectura (225)
 - Equivalencia con relación al ajuste personal-social (227)
 - Conclusiones sobre la equivalencia entre los grupos (228)
- 5.4 La muestra de la aplicación piloto (229)

6. Instrumentos /231

- 6.1 Instrumentos para las fases pretest y posttest (232)
 - El test de factor «G» de Cattell (232)
 - La escala APSL (233)
 - Validez y fiabilidad de la prueba APSL (238)
 - La aplicación piloto de APSL (238)
 - La validación de expertos (245)
 - Repercusiones de la validación de expertos para la versión final de la escala (249)
 - Las pruebas de comprensión de textos. La prueba ECO 1 y la prueba ECOS (250)
 - Análisis del contenido de la prueba ECO 1 (evaluación de la comprensión de textos) para segundo y tercer ciclo de educación primaria (252)
 - Dimensiones de la comprensión contempladas (254)
 - Validez y fiabilidad de la prueba ECO 1 (255)
 - Aplicación piloto (255)
 - La validación de expertos (262)
 - Repercusiones de la validación de expertos para la versión final de la prueba ECO 1 (266)
 - Análisis del contenido de la prueba ECOS (evaluación de la comprensión de textos) para educación secundaria (267)
 - Dimensiones de la comprensión contempladas (269)
 - Validez y fiabilidad en la prueba ECOS (270)
 - Aplicación piloto (270)
 - La validación de expertos (278)
 - Repercusiones de la validación de expertos para la versión final de la prueba ECOS (282)
 - Otros instrumentos: los cuestionarios EIDAL y EIDAL-MP sobre expectativas e ideas acerca de las dificultades de aprendizaje de la lectura (283)
 - Estructura y contenido de los cuestionarios EIDAL y EIDAL-MP (284)
- 6.2 El dispositivo de evaluación dinámica de procesos implicados en la lectura (EDPL) (286)
 - Lo que es y lo que no es el dispositivo EDPL (286)
 - Construcción del dispositivo y aplicación piloto (287)
 - Los procesos implicados en la comprensión de textos que han sido contemplados en el dispositivo (291)
 - Indicadores de evaluación de los procesos implicados (292)
 - El registro de la evaluación del proceso de resolución de las actividades y de los resultados obtenidos (297)
 - Esquema general de la Evaluación dinámica de procesos implicados en la lectura y conexiones entre las actividades de evaluación propuestas (298)
 - Esquema de las actividades de evaluación (298)
 - Esquema del desarrollo de las actividades de evaluación (300)
 - Galaxia de las actividades de evaluación (301)

7. Procedimiento /303

Resultados

8. Resultados en las pruebas de comprensión de textos, ajuste personal-social e inteligencia general /319

8.1 Las pruebas de comprensión de textos (319)

Resultados en la prueba de comprensión de textos ECO 1 para educación primaria (322)

Criterio de designación previa del profesorado sobre las dificultades de aprendizaje (322)

Contrastes en función de los distintos subgrupos considerados en la muestra (325)

Contrastes en función de los grupos de aplicación del dispositivo (327)

Criterio del profesorado + puntuación inicial en la prueba ECO 1 (331)

Criterio del profesorado + puntuación inicial en la prueba ECO 1 + puntuación inicial en Cattell (334)

Las DDP en la prueba de comprensión ECO 1 (336)

Resultados en la prueba de comprensión de textos ECOS para educación secundaria (339)

Criterio de designación previa del profesorado sobre las dificultades de aprendizaje (339)

Contrastes en función de los distintos subgrupos considerados en la muestra (341)

Contrastes en función de los grupos de aplicación del dispositivo (344)

Criterio del profesorado + puntuación inicial en la prueba ECOS (347)

Criterio del profesorado + puntuación inicial en la prueba ECOS + puntuación inicial en Cattell (349)

Las DDP en la prueba ECOS (350)

8.2 Resultados en la escala de ajuste personal-social APSL (352)

Criterio de designación previa del profesorado sobre las dificultades de aprendizaje (354)

Contrastes en función de los distintos subgrupos considerados en la muestra (359)

Contrastes en función de los grupos de aplicación del dispositivo (361)

Criterio del profesorado + puntuación inicial en la escala APSL (368)

Criterio del profesorado + puntuación inicial en la escala APSL + puntuación inicial en Cattell (370)

Relación entre el ajuste personal-social y la comprensión de textos (372)

Las DDP en la escala APSL (374)

Análisis de la fiabilidad de las mejoras en el grupo experimental (377)

8.3 Resultados en el test de inteligencia general de Cattell (380)

Criterio de designación previa del profesorado sobre las dificultades de aprendizaje (381)

Criterio del profesorado + puntuación inicial en las pruebas ECO 1 y ECOS (383)

Criterio del profesorado + puntuación inicial en las pruebas de comprensión + puntuación inicial en Cattell (384)

Las DDP en Cattell (385)

8.4 La Teoría de Respuesta al Ítem en las pruebas de comprensión de textos (387)

8.5 Los resultados de la muestra piloto en el postest (393)

Resultados postest de la muestra piloto en las pruebas de comprensión (394)

Resultados postest de la muestra piloto en la escala de ajuste personal-social (397)

9. Evaluación complementaria /401

- La evaluación complementaria: esquema del análisis propuesto (401)
- 9.1 La evaluación de la aplicación del dispositivo de evaluación dinámica de procesos implicados en la lectura (EDPL) (402)
 - Control experimental: análisis de las dificultades y vías de solución (403)
 - La cuantificación de las valoraciones en EDPL (406)
 - La matriz de datos en EDPL (412)
- Los resultados obtenidos en el dispositivo (414)
 - Evolución de las puntuaciones dinámicas en el transcurso de la aplicación de EDPL (422)
 - Relaciones entre la necesidad de mediación y la puntuación final obtenida (424)
 - Los resultados del dispositivo EDPL en función de los grupos de aplicación (425)
- Análisis del dispositivo en función del perfil de cuatro puntos (426)
 - Información comparativa del dispositivo EDPL (426)
 - Calidad de los datos obtenidos. La fiabilidad de las valoraciones en EDPL (427)
 - El tipo de información ofrecida y las aportaciones metodológicas adicionales (430)
 - El valor predictivo de EDPL (431)
 - Valor de predicción con relación al rendimiento académico (433)
 - Valor predictivo en la prueba de comprensión ECO 1 (436)
 - Análisis de regresión en la prueba ECO 1 (438)
 - El análisis de regresión para los grupos de mejoradores y no mejoradores (440)
 - Valor predictivo en la prueba de comprensión ECOS (444)
 - Análisis de regresión en la prueba ECOS (447)
 - El análisis de regresión para los grupos de mejoradores y no mejoradores (449)
 - Valor de predicción con relación a la escala APSL (453)
 - Análisis de regresión en la escala APSL (457)
 - Valor de predicción con relación a la prueba de CI (459)
 - Análisis de regresión en la prueba de CI (462)
 - Grado de eficacia (464)
 - Análisis de la información cualitativa recogida por los aplicadores del dispositivo EDPL (465)
 - El análisis cualitativo del contenido de las Hojas de registro y evaluación (465)
 - Tipo de información recogida y destacada por los aplicadores (466)
 - Relación entre el tipo de información recogida y destacada por los aplicadores, y las puntuaciones dinámicas (475)
 - Patrones metodológicos de evaluación dinámica de la lectura utilizados en la mediación (477)
 - Estructura factorial (494)
 - Análisis y valoración conjunta del Grupo de Trabajo y del dispositivo de evaluación dinámica EDPL (497)
 - Análisis y valoración conjunta de la labor desarrollada en el Grupo de Trabajo (498)
 - Valoración del dispositivo de evaluación dinámica de procesos implicados en la lectura (EDPL) y de su aplicación (503)
 - Enriquecimiento de la propuesta mediante la ampliación de las actividades (506)
 - Algunos ejemplos de la mediación efectuada por los aplicadores (507)
 - 9.2 Análisis de los cuestionarios EIDAL y EIDAL-MP sobre expectativas e ideas acerca de las dificultades de aprendizaje de la lectura (508)
 - Análisis exploratorio del cuestionario EIDAL (508)
 - Relación entre las puntuaciones de mejora y las respuestas en EIDAL (512)
 - Análisis exploratorio del cuestionario EIDAL-MP (519)
 - Relación entre las puntuaciones de mejora y las respuestas en EIDAL-MP (523)

Discusión y Conclusiones

10. Discusión /533

- 10.1 Discusión general relativa a las bases teóricas y metodológicas del dispositivo de evaluación dinámica EDPL (533)
- 10.2 Discusión sobre el proceso de aplicación experimental de EDPL (541)
 - Estructura, contenidos e información analizada en el dispositivo EDPL (541)
 - Conexión de la aplicación experimental con nuestra teoría de base y mecanismos de acción del dispositivo EDPL (548)
- 10.3 Discusión del proceso de evaluación seguido para valorar la aplicación experimental del dispositivo EDPL: metaevaluación (551)
- 10.4 Discusión de los resultados en las pruebas de la batería criterio (558)
 - Discusión de los resultados en la prueba de comprensión de textos para educación primaria ECO 1 (558)
 - Discusión de los resultados en la prueba de comprensión de textos para educación secundaria ECOS (562)
 - Discusión de los resultados en la escala de ajuste personal-social APSL (565)
 - Discusión de los resultados en el test de inteligencia general de Cattell (572)
 - Discusión de los resultados de la muestra piloto en el postest y de los análisis realizados en función de la Teoría de Respuesta al Ítem (574)
- 10.5 Discusión de los resultados en la evaluación complementaria (577)
 - Discusión del análisis del dispositivo en función del perfil de cuatro puntos (581)
 - Discusión de los resultados del análisis de los cuestionarios EIDAL y EIDAL-MP (597)
 - Discusión sobre la incidencia de la aplicación del dispositivo de evaluación dinámica EDPL en la mejora del proceso de evaluación (601)
 - Discusión sobre las propuestas de mejora del dispositivo EDPL (602)

11. Conclusiones /605

- Con relación a la comprensión de textos (605)
- Con relación al ajuste personal-social (608)
- Con relación a la inteligencia general (611)
- Con relación a la mejora del proceso de evaluación y al proceso de aplicación (611)
- Con relación a los procesos metacognitivos, su relevancia y su dificultad (617)
- Con relación a la aplicación piloto (619)
- Con relación al valor explicativo diferencial de EDPL sobre el rendimiento (619)

Referencias Bibliográficas /623

Anexos (*Volumen de Anexos*)

- Anexo I.* Prueba ECO 1. Evaluación de la comprensión de textos (E. primaria)
- Anexo I-A* Claves para la corrección de la prueba ECO 1
- Anexo II.* Prueba ECOS. Evaluación de la comprensión de textos (E. secundaria)
- Anexo II-A* Claves para la corrección de la prueba ECOS
- Anexo III.* Escala APSL. Escala de ajuste personal-social con relación a la lectura
- Anexo III-A* Claves de corrección de las escalas APSL 1 y APSL
- Anexo IV.* Hoja de registro y valoración de la escala APSL
- Anexo V.* Plantilla de datos relacionados con la investigación
- Anexo VI.* Plantilla de validación de expertos (prueba ECO 1)
- Anexo VII.* Plantilla de validación de expertos (prueba ECOS)
- Anexo VIII.* Plantilla de validación de expertos (escala APSL)
- Anexo IX.* Plantilla de valoración del profesorado acerca del rendimiento escolar en el área de Lengua
- Anexo X.* Valoración del dispositivo de Evaluación Dinámica de Procesos implicados en la Lectura (EDPL) y de su aplicación
- Anexo XI.* Cuestionario EIDAL
- Anexo XI-A.* Análisis del cuestionario EIDAL. Sistema de categorías elaborado y suma de las referencias expresadas por el profesorado
- Anexo XII.* Cuestionario EIDAL-MP
- Anexo XII-A.* Análisis del cuestionario EIDAL-MP. Sistema de categorías elaborado y suma de las referencias expresadas por el profesorado
- Anexo XIII.* Actividades recogidas en el dispositivo EDPL que han sido objeto de análisis complementario
- Anexo XIV.* Cuaderno de Hojas de registro y evaluación
- Anexo XV.* Revisión bibliográfica con relación a la evaluación de procesos implicados en la comprensión
- Anexo XVI.* Propuesta de actividades de evaluación dinámica de los distintos procesos contemplados en EDPL, y de intervención centrada en los procesos metacognitivos
- Anexo XVII.* Enriquecimiento de la propuesta mediante la ampliación de las actividades
- Anexo XVIII.* Algunos ejemplos de la mediación efectuada por los aplicadores (transcripción y análisis de la mediación efectuada)

Introducción

La evaluación dinámica es uno de esos conceptos que ha cautivado el interés de un buen número de psicólogos a lo largo de varias décadas. Atravesando la psicología con vocación continua de ampliar los límites establecidos en cada una de sus disciplinas, estableciendo puentes entre distintos ámbitos de conocimiento y de trabajo práctico (como la psicología educativa, la pedagogía y la didáctica), o tratando de conjugar conocimiento psicológico y práctica educativa, la evaluación dinámica ha suscitado, sin embargo, al menos tantas dudas como certezas si nos atenemos a los resultados de la investigación aplicada (Mora, 1997; Grigorenko y Sternberg, 1998; Sternberg y Grigorenko, 2003).

Por otra parte, la historia de la evaluación dinámica, desde que en las primeras décadas del siglo veinte, Vygotski (1934/1995) propusiera por primera vez la necesidad y la conveniencia de evaluar de forma interactiva el proceso de aprendizaje de un sujeto, ha estado caracterizada por los continuos intentos por parte de algunos investigadores de «aprisionar» dentro de los a veces estrechos márgenes de la psicología psicométrica, los postulados teóricos y prácticos que se habían generado como alternativa a una evaluación de corte estático y convencional, que mostraba sus limitaciones especialmente en el ámbito educativo. En efecto, la tradición psicométrica en evaluación educativa no podía dar cuenta de los procesos de cambio y aprendizaje producidos como consecuencia de la acción educativa, dificultando significativamente de esta forma el conocimiento sobre los procesos afectados y sobre el modo en que las dificultades podrían ser reducidas (Mora, 1997; Coll, Barberà y Onrubia, 2000; Coll, Martín y Onrubia, 2001). Paradójicamente, la aplicación de este mismo paradigma en la investigación con pruebas de evaluación dinámica —lo que conlleva en la práctica el uso de técnicas psicométricas para valorar la aplicación de las pruebas diseñadas bajo modelos dinámicos—, podría estar teniendo una doble consecuencia: por una parte, podría estar frenando y encorsetando el desarrollo de estas pruebas, ya que, en ocasiones, los modelos propuestos, aún pudiendo ser de interés desde una perspectiva práctica y centrada en la mejora de los procesos educativos, no se dejan analizar fácilmente desde enfoques exclusivamente psicométricos y que no contemplan más allá de resultados cuantificables; por otra, en cambio, podría estar contribuyendo a ofrecer un marco común para la interpretación de los resultados, permitiendo su generalización, así como la de los propios procesos de evaluación dinámica con el objetivo de que puedan ser realmente útiles para la optimización de los procesos de enseñanza y aprendizaje. Hemos de decir que en este difícil equilibrio se mueve la investigación que se presenta en este trabajo.

La evaluación dinámica se apoya en la interacción del sujeto evaluado con otro más experto, que presta su ayuda estableciendo un proceso de mediación entre la capacidad actual del sujeto y las exigencias de la tarea; perfilando de este modo un espacio de construcción del conocimiento sensible a la acción pedagógica —lo que constituye la *zona de desarrollo próximo o potencial*— (Wertsch, 1988; Newman, Griffin y Cole, 1989; Newman y Holzman, 1993; Rogoff, 1993; Jensen, 2000); todo esto con objeto de determinar qué tipo de ayuda educativa es necesaria, y estableciendo una medida dinámica de los procesos cognitivos que se encuentran en continuo desarrollo, en

contraposición a una medida estática de un constructo psicológico que informe sobre las habilidades cognitivas ya consolidadas por el sujeto. La relevancia de la aplicación de modelos dinámicos de evaluación parece ser mayor si cabe ante sujetos que presentan dificultades en su aprendizaje o discapacidad (Campione y Brown, 1987; Carlson y Wiedl, 2000; Guthke y Beckmann, 2000; Swanson, 2000; Saldaña, 2001). Por otra parte, la consideración, en los procesos de evaluación, del contexto social caracterizado por la interacción y el apoyo en el que se desenvuelve habitualmente el proceso de enseñanza y aprendizaje escolar, puede contribuir significativamente al incremento de la validez ecológica de estos procedimientos, facilitando la replicación de los procesos aplicados y la obtención de resultados más generalizables (Schlee, 1985; Mora, 1997; Navarro, 2004).

A pesar de que Vygotski hizo referencia a la evaluación dinámica en procesos de valoración tanto de aptitudes cognitivas como de otras más ligadas a los procesos de aprendizaje escolar, el interés primordial de los investigadores ha estado dirigido hacia la aplicación de las pruebas dinámicas para valorar las primeras; en concreto, la mayoría de las pruebas elaboradas han tenido por objeto la evaluación de las habilidades relacionadas habitualmente con la inteligencia (Kozulin y Garb, 2001). Sin embargo, en los últimos años se ha suscitado un renovado interés por la evaluación dinámica de procesos implicados en el aprendizaje de contenidos escolares. Este interés ha tenido su origen en la constatación de la relevancia de varios factores: (a) la especificidad de contenido de las pruebas en la explicación de los resultados y de las mejoras observadas; (b) la contextualización de los contenidos de las pruebas con relación al espacio que se pretende valorar y en el que se pretenden introducir mejoras; y (c) la generalización de los procesos de aplicación de las actividades desarrolladas y de los resultados obtenidos. Esto ha contribuido de manera significativa a que se produzca un replanteamiento de la evaluación dinámica y de sus posibles aportaciones al proceso de enseñanza y aprendizaje escolar (Sternberg y Grigorenko, 2003; Navarro, 2004).

Uno de esos contenidos escolares a los que hacíamos referencia anteriormente es la lectura. En efecto, la preocupación y el interés creciente que los procesos de enseñanza y aprendizaje de la lectura han originado en la psicología y en la educación en las últimas décadas, ha desembocado en la aplicación progresiva de nuevas o renovadas perspectivas teóricas y prácticas, derivadas a su vez de la investigación psicológica, y aplicadas, tanto a la intervención educativa para la mejora del proceso de lectura y de las estrategias de comprensión, como a los procesos de evaluación (Alonso Tapia y Cols., 1992; García Sánchez, 2001; Navarro y Mora, 2003a). Entre estas renovadas perspectivas desde las que abordar los procesos educativos relacionados con la evaluación de la lectura, originadas no pocas veces en la constatación de las limitaciones que ofrece la aplicación convencional de pruebas estandarizadas centradas en el producto de la comprensión, se encuentran los procedimientos dinámicos de evaluación; si bien es verdad que son aún muy escasos los estudios que se han adentrado en la evaluación dinámica de los procesos implicados en la lectura, manteniendo además éstos entre sí una gran diversidad en cuanto al ámbito de aplicación, la metodología empleada, los contenidos específicos evaluados, así como en cuanto a sus objetivos.

Otro de los aspectos relacionados con la lectura que ha experimentado una cierta revolución en su conceptualización y en las aplicaciones prácticas a que ha dado lugar, ha sido precisamente el contenido de lo que debe ser enseñado y evaluado. En este

sentido, los descubrimientos y redescubrimientos con relación a los procesos que están implicados en la comprensión del lenguaje escrito, han ido poniendo de manifiesto la relevancia del control y regulación consciente del proceso de lectura, así como del conocimiento consciente de la estructura del lenguaje escrito o de las estrategias de lectura que pueden emplearse para lograr la comprensión. Igualmente han tomado una especial importancia los procesos relacionados con el contexto educativo y social en el que se desarrolla el aprendizaje y la enseñanza escolar, así como los aspectos relativos a la motivación y las actitudes (Alonso Tapia y Cols., 1992; Solè, 1992; Sánchez, 1999; Montanero, 2000; Moreno y García, 2000; Kozulin y Garb, 2001; García Sánchez, 2002; Bofarull, 2001; Navarro y Mora, 2002; Romero y Lavigne, 2006). El análisis preferente de los procesos metacognitivos implicados en la comprensión, mediante la aplicación de un modelo dinámico de evaluación, como medio privilegiado para acceder a las posibles dificultades de aprendizaje de la lectura y al modo en que podría optimizarse el proceso de lectura, constituye el eje central de nuestro estudio y también de nuestro análisis. Dedicaremos los primeros capítulos de este trabajo, antes de adentrarnos en el desarrollo de la investigación, precisamente a revisar y seguir componiendo el marco teórico relacionado con la lectura, los procesos implicados en la comprensión y los modelos dinámicos de evaluación.

Por otra parte, *lectura y comprensión* son dos conceptos que utilizaremos en ocasiones indistintamente a lo largo del presente trabajo, ya que, aunque somos conscientes de que cuando se hace referencia a la lectura pueden quedar incluidos habitualmente una serie de procesos, como la decodificación, que no suelen aparecer al abordar los procesos de comprensión de textos, consideramos la lectura fundamentalmente como un proceso encaminado a la comprensión (Solè, 1992; Solè y Teberosky, 2001; Repetto, Téllez y Beltrán, 2002; Navarro y Mora, 2002), que alcanza su sentido y significado en ésta, conjugando a su vez distintos procesos que interactúan con el objetivo de lograr una representación compleja y global de los contenidos que se dan cita cuando se lee un texto.

MARCO TEÓRICO

1. Naturaleza y evaluación de la lectura

1.1 *La lectura como construcción socio-histórica y dinámica*

La lectura es un proceso que ha sido construido socialmente en contextos de interacción. Los diversos análisis realizados desde un enfoque socio-histórico (Solè, 2001), apuntan la posibilidad de que los intentos del ser humano por plasmar el lenguaje en soportes de lo más variados y con fines también diversos, que iban desde el saber popular hasta las normas de convivencia o las creencias; se transformaran progresivamente en la convención social de signos gráficos que pudieran ser descodificados e interpretados por los miembros de la cultura de referencia. De igual forma, la lectura es un proceso que se construye socialmente en un contexto de interacción cuando hablamos de leer y comprender un texto. Incluso cuando quien afronta la lectura es una sola persona, se trata igualmente de un proceso interactivo (Solè, 1992, 2001; Alonso Tapia y Cols., 1992; Sánchez, 1999; Solè y Teberosky, 2001; Moreno y García, 2000) entre las características formales, contextuales y de contenido del propio texto, y las características del sujeto que lee, sus habilidades, intereses, conocimientos, motivación y actitudes.

Por otra parte, la constatación de que la estructura misma del lenguaje —como actividad psicológica— se ha originado alrededor de un núcleo pragmático, es decir: la progresiva construcción y convención social de signos lingüísticos; y de que el lenguaje en sí, según Vygotski (1934/1995), se originaría continuamente a partir de las funciones comunicativas y de interacción; no hace sino reforzar y confirmar el carácter social y de interactividad del proceso. Esta premisa es válida también con relación al lenguaje escrito, que, como apuntábamos más arriba, es construido —históricamente y cada vez que se actualiza el hecho de leer o escribir— en continua interacción. Podríamos decir que cuando se escribe, se hace para comunicar algo, y cuando se lee, se hace para entender y conocer algo que se ha querido comunicar.

Una consecuencia relevante de las ideas expresadas en los párrafos anteriores, apunta a que todos los elementos que conforman el proceso de lectura, se constituyen a su vez en procesos interactivos; tanto aquellos más relacionados con las conductas cognitivas que debe poner en práctica el sujeto para acceder al código y a la comprensión del texto, como también aquellos relacionados con los factores contextuales que favorecen o dificultan el acceso a la lectura; y son interactivos, no sólo porque un análisis de sus contenidos desvelaría una continua inter-relación de factores, sino porque la interacción de estos procesos *permite* el hecho de leer. Si esto es así, de la misma forma, aquellos elementos característicos de la educación escolar, como la evaluación del proceso de enseñanza y aprendizaje de la lectura y la intervención educativa para favorecer el proceso de adquisición de habilidades lectoras o tratar de solucionar problemas en el dominio de la lectura, deben considerarse procesos que se construyen en interacción, y la puesta en práctica de tareas, pruebas o actividades encaminadas a evaluar o intervenir sobre estos aspectos, debe contar con esto.

La lectura es también, además de un proceso interactivo, un proceso dinámico. La naturaleza de la lectura como actividad social y cognitiva revela en esencia su potencial

de transformación continua, tanto de estructuras relativas a los contenidos del pensamiento, como a la destreza en el acto mismo de leer y comprender los textos. La modificabilidad funcional o estructural del proceso de lectura podría ser contemplada como consustancial a su aprendizaje y, por tanto, a los intentos de su mejora cuando existen dificultades.

La concepción de la lectura que va a servir de marco al desarrollo de este trabajo y en la que, por consiguiente, se encuadran y toman sentido tanto las propuestas teóricas y metodológicas, como las decisiones adoptadas en el estudio presentado, se caracteriza por considerar la lectura como:

- a) un proceso interactivo (Solè, 1992; Solè y Teberosky, 2001) y dinámico;
- b) una construcción social que tiene lugar en contextos de interacción (Vygotski, 1934/1995; Luria, 1995; Navarro y Mora, 2002);
- c) un proceso de construcción y representación global de significados en el que tienen lugar a su vez procesos cognitivos y metacognitivos, motivacionales, emocionales, lingüísticos y socio-culturales (Repetto, Téllez y Beltrán, 2002; Solè, 1992; Navarro y Mora, 2002; Romero y González, 2001);
- d) un proceso activo y susceptible de ser llevado a la conciencia, de ser controlado y regulado, y de ser objeto de reflexión (Palinskar y Brown, 1984; Johnston, 1989; Repetto, Téllez y Beltrán, 2002; Navarro y Mora, 2003a; Sánchez, 1999; Alonso Tapia et al., 1992; Alonso y Carriedo, 1996; Mateos, 2001);
- e) un proceso dialéctico y de reflexión, que permite la trascendencia de lo inmediato, está al servicio del aprendizaje y la comprensión, el pensamiento y el análisis crítico, la imaginación, las emociones y la transformación social (Mayor Zaragoza, 2002; Elboj, Puigdemívol, Soler y Valls, 2002; Navarro, 2005).

Desde esta concepción de la lectura como construcción dinámica y social en contextos de interacción, se asume que la lectura está estrechamente relacionada con el pensamiento, con la construcción y reconstrucción de actitudes, ideas y procedimientos, con el aprendizaje y la comprensión, con la transformación social y personal, con las emociones y los sentimientos, la imaginación y el descubrimiento.

Abordaje conceptual de la lectura: características básicas de los modelos explicativos

El abordaje conceptual de la lectura ha tenido lugar históricamente desde tres modelos explicativos distintos. Estos se han ido desarrollando en el tiempo de forma que el primero de ellos, que encuadra a su vez a una serie de teorías explicativas no siempre coincidentes plenamente, conocido como *modelos ascendentes*, fue dando paso al segundo a raíz del análisis de su práctica y de la constatación de sus debilidades para dar cuenta de una forma más completa de cómo tenía lugar el proceso de lectura; así mismo, fue la constatación de los problemas que este segundo grupo de modelos explicativos, los llamados *modelos descendentes*, presentaba para explicar y guiar por sí solo el concepto y el aprendizaje de la lectura, lo que dio origen al tercer grupo de modelos, los *modelos interactivos*. En definitiva, tanto las concepciones sobre la lectura como, de forma paralela, los modelos explicativos de su aprendizaje, se han ido sucediendo a medida que las investigaciones y la práctica educativa iban revelando ciertas debilidades o inconsistencias de sus propuestas. Aunque el contenido de estos modelos de lectura es bien conocido, nos conviene hacer una breve descripción de cada uno de ellos, simulando la secuencia histórica que han seguido en su aparición en el

ámbito de la psicología educativa, para detenernos finalmente en los modelos interactivos, que representan en la actualidad el marco conceptual y metodológico en el que se desarrollan la gran mayoría de las propuestas educativas relacionadas con la lectura.

Los *modelos ascendentes* (Gough, 1976; Laberge y Samuels, 1976), o también llamados *modelos regidos por datos*, se caracterizan básicamente por enfatizar los procesos de descodificación de las unidades componentes del lenguaje. Este proceso de descodificación conduciría de forma automática a la comprensión, la cual tendría lugar a través de la conjunción entre el reconocimiento visual de las palabras y la comprensión oral. El proceso descrito en las propuestas que se encuadran en estos modelos consiste en una secuencia ascendente y jerárquica desde las unidades lingüísticas más pequeñas y sin significado (letras, fonemas, sílabas) hasta las unidades mayores ya con significado (palabras, oraciones, textos...).

La consideración crítica de los modelos ascendentes de aproximación a la lectura y su aprendizaje es de enorme importancia, ya que éstos conllevan a menudo una determinada forma de entender cómo se produce nuestro aprendizaje y cómo tiene lugar el pensamiento. Del mismo modo, estos modelos también generan un modo de entender la mejor forma de intervenir ante las dificultades en el desarrollo del proceso lector (Solè, 2001). De esta forma, tal como destacan Repetto, Téllez y Beltrán (2002), los estudios y propuestas enraizados esencialmente en estos modelos han acentuado la importancia de intervenir con métodos reeducativos de problemas de lecto-escritura basados en *procesos perceptivos y atencionales* (velocidad lectora, discriminación perceptiva y auditiva) y de *memoria* (reconocimiento de patrones y de vocabulario); así mismo sus tesis teóricas se han relacionado con los métodos de enseñanza de la lectura denominados *sintéticos*.

La constatación de que en la comprensión influían de forma significativa variables tales como los conocimientos previos del lector, su conocimiento sobre la estructura textual y la sintaxis, dio lugar a nuevos planteamientos teóricos y prácticos que podrían encuadrarse en el siguiente grupo de modelos explicativos que revisaremos.

Los *modelos descendentes* (Goodman, 1967, 1982a, 1982b; Smith, 1983), o *modelos regidos por conceptos*, destacan la relevancia que para el proceso de comprensión tiene el contexto en el cual se produce la lectura, es decir, reconocer las claves textuales que permiten al lector generar determinadas hipótesis y predicciones que pueden guiar y facilitar su comprensión, de forma que ésta tiene lugar por la anticipación de los contenidos de un texto con base en los conocimientos previos, relativos al contenido o a la estructura textual, que posea el sujeto. En este caso, se describe también una cierta secuencia jerárquica, aunque descendente, desde unidades globales de tipo conceptual e incluso extra-lingüístico (los conocimientos previos), hasta las unidades lingüísticas que conforman la base textual.

Las propuestas a que han dado lugar los trabajos enmarcados en estos modelos enfatizan la activación de conocimientos previos para facilitar y guiar la comprensión, el entrenamiento para el conocimiento sintáctico, que permite también realizar predicciones, y el conocimiento de las distintas tipologías de textos para facilitar la aplicación de esquemas interpretativos con base en la estructura (Repetto, Téllez y Beltrán, 2002). El reconocimiento de la relevancia del contexto y de los conocimientos

previos del lector, esgrimida en los modelos descendentes, no hizo olvidar, sin embargo, la necesidad de que existiera, de una u otra forma, algún tipo de procesamiento perceptivo con base en el texto que se debía leer. Esta realidad fue la que dio origen a postulados teóricos y propuestas metodológicas de corte más interactivo.

Los *modelos interactivos* (Rumelhart, 1976, 1977, 1980; Rumelhart y McClelland, 1981; Kintsch y Van Dijk, 1978; Kintsch, 1986; Ruddell y Spieker, 1985; Ruddell y Ruddell, 1994; Adams y Collins, 1977; Adams, 1994; Just y Carpenter, 1987) se caracterizan, tal y como hemos destacado desde el comienzo de este trabajo, por entender la lectura como un proceso en el que actúan de forma conjunta e interrelacionada diversos elementos, cuya interacción explica y permite la lectura, que a su vez es entendida como un proceso de construcción y representación global de significados. Esta interacción puede tener lugar a distintos niveles: (a) entre el lector y el texto, (b) entre procesos centrados en la decodificación y aquellos más relacionados con la continua generación de hipótesis y predicciones sobre la base de los conocimientos previos, (c) entre las características del texto, las del lector y las del contexto socio-cultural.

De esta forma, el modelo interactivo integra, pero también trasciende las aportaciones realizadas anteriormente por las aproximaciones centradas en los procesos de corte ascendente o descendente (Solè, 1992, 2001). Las propuestas e investigaciones realizadas desde este modelo —la gran mayoría en los últimos años— asumen básicamente las características citadas anteriormente, aunque desde diversos enfoques interpretativos, abordando como consecuencia también de forma distinta la enseñanza y aprendizaje de la lectura. A pesar de que todos parten de supuestos teóricos en los que la interacción de los procesos y elementos implicados en la lectura —especialmente entre el lector y el texto— es el componente relevante, existen notables diferencias en la práctica derivadas fundamentalmente de la consideración del proceso de comprensión como un conjunto de subhabilidades que, a su vez, pueden ser entrenadas de forma más o menos independiente, o como un proceso unitario. En línea con la primera de estas ideas se encontrarían los trabajos de Baumann (1985, 1986) y su propuesta de enseñanza directa de las diferentes habilidades de comprensión lectora. Baumann estructura su propuesta básicamente en cuatro fases: (a) *explicación verbal*, en la que se introducen y se ejemplifican aquellas habilidades que se pretenden enseñar; (b) *modelado verbal*, en la que el profesor o profesora muestra y describe la habilidad en voz alta mientras realiza la tarea; (c) *práctica guiada* o aplicación dirigida por el profesor, en la que el alumnado practica la habilidad bajo la supervisión del docente; y (d) *práctica independiente*. Solè (2001) hace referencia a las limitaciones de este enfoque aludiendo a que, inspirándose en un modelo de investigación ‘proceso/producto’, asume una cierta causalidad y linealidad entre la actividad ejercida por el docente y el resultado final, es decir, en este caso, entre la puesta en práctica rigurosa de las distintas fases descritas y el aprendizaje del alumno. Según esta autora se vería de este modo comprometida la actividad mental constructiva del propio alumno o, lo que es lo mismo, su participación activa como parte del proceso interactivo en la dotación de sentido y significado a su lectura. Por otra parte, también podría entrañar un cierto riesgo descomponer en subhabilidades menores lo que pueden ser estrategias más complejas. La presunción de que estas subhabilidades están vinculadas y forman parte de un todo, así como que el dominio de éstas puede posteriormente ser generalizado e integrado ante actividades que requieran una lectura contextualizada y una comprensión global, no parece hallarse suficientemente sustentada.

La segunda de estas consideraciones de la lectura a la que hacíamos mención dentro de los modelos interactivos, hace referencia a la comprensión como un proceso unitario. La comprensión se aborda desde una óptica más centrada en la capacidad del sujeto para controlar y regular el proceso de lectura y las propuestas metodológicas van encaminadas a enseñar estrategias que permitan interpretar los textos de forma progresivamente autónoma, autointerrogarse sobre su contenido, planificar la propia actividad, supervisar y controlar si se van alcanzando los objetivos propuestos, o evaluar todo el proceso y la propia comprensión. Se trata de propuestas, por tanto, que acentúan la actividad constructiva y metacognitiva del lector. El conocido modelo de *Enseñanza Recíproca* de Palinskar y Brown (1984), la más reciente propuesta de *Intervención Estratégica en Niveles (IEN)* de Montanero, Blázquez y León (2002), los programas *Mejorar nuestra comprensión y expresión escrita* (Hernández y Quintero (2001) o *Comprender y Aprender en el aula* (Repetto, Sutil, Manzano, Téllez y Beltrán, 2002) y en definitiva, todas aquellas propuestas, integradas o no en programas elaborados y sistemáticos, que inciden en la actividad primordial del alumno en el proceso de construcción de la comprensión, en la consideración de la presencia conjunta de las distintas estrategias en el desarrollo de la actividad lectora y en el énfasis en una ayuda educativa dirigida especialmente a favorecer un desarrollo progresivamente autónomo del proceso; constituyen algunos de los ejemplos de esta perspectiva.

Por su parte Rosenblatt (1978), en su *teoría transaccional* sobre la enseñanza de la literatura destaca de una forma especial el papel activo del lector en la lectura. Utiliza el término transacción del lector con el texto para hacer referencia a la configuración de una relación especial entre ambos elementos, introduciendo un matiz distinto con respecto al de interacción; pone un mayor énfasis en lo indisoluble de la relación entre el lector y el texto; el significado, de esta forma, será construido sólo mediante la conjunción, única y particular, del sujeto que lee con el texto. Actividades como la discusión, el diálogo y el debate en torno a los textos facilitarían, según esta autora, esta relación transaccional (Dubois, 2003). En general, podríamos decir que los modelos interactivos de lectura, especialmente aquellos que proponen acciones encaminadas a potenciar un uso autónomo del lenguaje escrito, entendiendo el proceso de enseñanza y aprendizaje desde una concepción constructivista, asumen e integran los elementos destacados por Rosenblatt, incluyendo el aspecto esencial de la teoría transaccional, es decir, la capacidad del lector para influir o modificar el proceso en el que interviene, en este caso la lectura.

Estrategias y procesos contemplados en la lectura

Por otra parte, el aprendizaje de la lectura como proceso construido en interacción y que sigue patrones de procesamiento caracterizados, a su vez, por la interacción de los elementos implicados, supone la puesta en práctica de una serie de estrategias que permitan al lector ser capaz de acceder a la comprensión global y profunda del contenido de los textos y aprender de éstos. Desde esta perspectiva de la lectura como proceso interactivo, se han descrito algunas de estas estrategias:

* Ser conscientes de qué significa leer y para qué sirve (Alonso Tapia y Carriedo, 1996; Maruny, Mistral y Miralles, 2001; Martín, 1999; Moreno y García, 2000; Repetto, 1997; Blanton, Wood y Moorman, 1990). Requiere saber qué es lo que debe conocerse y lo que se puede hacer para conseguir el objetivo marcado.

* El uso de claves textuales y del conocimiento previo para ser capaz de construir inferencias a raíz de información implícita contenida en el texto, integrar los nuevos contenidos con experiencias y conocimientos que ya se poseen y ser capaz de revisar y valorar críticamente la información nueva con base en lo ya conocido (Alonso Tapia y Carriedo, 1996; Sánchez, 1999; Vidal-Abarca y Gilabert, 1991; Hernández y Quintero, 2001; Meyer, 1984; 1985; González y Romero, 2001); estructurar y organizar la información y realizar predicciones sobre el significado de palabras, frases, fragmentos del texto, etc., pudiendo evaluar y revisar las hipótesis formuladas.

* El uso del contexto, tanto del contexto discursivo contenido en el texto, como del contexto más amplio que incluye la colaboración de profesores y compañeros, y el reconocimiento de situaciones sociales que influyen en el significado y sentido de un texto determinado en un momento concreto (Solè, 1992; Solè y Teberosky, 2001; Moreno y García, 2000; Edwards y Mercer, 1988; Mateos, 1991; Alonso y Carriedo, 1996).

* Estrategias de control y regulación de la propia comprensión y de sus estados motivacionales. Estas estrategias harían referencia al establecimiento de metas y objetivos en nuestra lectura, la secuenciación y jerarquización de operaciones a realizar, la supervisión del propio proceso de comprensión durante la lectura, dándonos cuenta de cuándo no estamos comprendiendo, la evaluación de todo el proceso y del resultado obtenido, valorando las estrategias utilizadas. También hacen referencia al control metacognitivo sobre cómo nos encontramos a la hora de afrontar la lectura y a cuál es nuestra actitud, interés y motivación para leer (Jolibert, 1995; Mateos, 1991; Alonso Tapia y Cols., 1992; Sánchez, 1998; 1999; Lin y Zabrusky, 1998; Moreno y García, 2000; Repetto, 1997; Hacker, 1998; Romero y González, 2001).

Estas estrategias mencionadas han sido extraídas, lógicamente, de procesos de auto-observación y observación de situaciones de aprendizaje en las cuales los sujetos han dado muestra de ponerlas en práctica y, como consecuencia, de comprender y aprender de los textos escritos. En el ámbito de estudio de la comprensión lectora ha existido un notable interés por determinar qué hacían bien los alumnos que manifiestan buen nivel de comprensión y qué es lo que hacían mal los alumnos que manifestaban problemas para acceder a la comprensión de lo escrito. Este análisis ha permitido configurar una serie de estrategias, como las anteriormente expuestas, que son consideradas idóneas, ya que su uso se ha relacionado con la comprensión y el aprendizaje a partir de los textos, y, como consecuencia, *deseables* para todo el alumnado. En otras palabras, se ha tratado de determinar cuáles son las estrategias que utilizan habitualmente los alumnos que leen bien, y pueden aprender de lo que leen y, al mismo tiempo, cuáles son las «estrategias» que utilizan los lectores con problemas en la comprensión y por qué. En un cuadro de Moreno y García (2000) que recogemos en la página siguiente, al que pusimos referencias sobre trabajos que apoyan los elementos citados, resumimos estos puntos.

Por otra parte, de forma reiterada se ha puesto de manifiesto que la lectura, así como su enseñanza y aprendizaje, es una actividad no exenta de complejidad (Sánchez, 1990; Solè, 1992), y no seremos nosotros quienes lo pongamos en duda. Basta tratar de profundizar un poco en los elementos y mecanismos necesarios en su realización, para desechar cualquier intento de explicación que corra el peligro de ser reduccionista o excesivamente simple. Ciertamente leer exige poner en funcionamiento de forma

Cuadro 1. Diferencias en la utilización de estrategias de comprensión (adaptado de Moreno y García, 2000)

<i>Qué hacen bien los alumnos/as con buen nivel de comprensión</i>	<i>Qué hacen mal los alumnos/as que presentan dificultades en la comprensión</i>
<ul style="list-style-type: none"> • Focalizan de una forma flexible y estratégica su lectura (Cooper, 1990; Hernández y Quintero, 2001; Sánchez, 1998; 1999). • Organizan la información (Just y Carpenter, 1987; Solè, 1992). • Gestionan estrategias para solucionar problemas de comprensión (Palinskar y Brown, 1984; Solè, 1992; Brown, 1980; Montanero, Blázquez y León, 2002). • Elaboran la información textual y la integran con sus conocimientos y con otras fuentes de información (generalización) (Palinskar y Brown, 1984; Solè, 1992; Carriedo, 1996; Sánchez, 1998). • Verifican la información obtenida o supuesta (Solè, 1992; Buron, 1993; Jolibert, 1995; Repetto, Téllez y Beltrán, 2002). • Son conscientes de los fines de la lectura (Palinskar y Brown, 1984; Solè, 1992; Blanton, Wood y Moorman, 1990; Repetto, Téllez y Beltrán, 2002). • Entienden la lectura como un proceso (Solè, 1992). • Tienen mayor metaconocimiento estratégico (Alonso Tapia y Col., 1992; Montanero, 2001; Navarro y Mora, 2003a). • Uso estratégico de la progresión temática del texto (Sánchez, 1993; 1999) • Uso de estrategias estructurales que permiten establecer relaciones significativas entre los distintos valores textuales (Vidal-Abarca y Gilabert, 1991; Sánchez, 1993; 1998; 1999; Meyer, 1984; 1985; Meyer, Brandt y Blouth, 1980; Van Dijk, 1983). • Buena autorregulación del proceso lector (Solè, 1992; Sánchez, 1998; 1999; Repetto, 1997; Vidal-Abarca y Gilabert, 1991; Mateos, 1991; Jolibert, 1995; González, 1996; 1997; Lin y Zabusky, 1998; Hacker, 1998; Repetto, Téllez y Beltrán, 2002). 	<ul style="list-style-type: none"> • No focalizan de forma estratégica su atención (Palinskar y Brown, 1984). • No organizan la información, sino que abordan el proceso de forma lineal (Sánchez, 1999; González y Romero, 2001). • No detectan sus errores de comprensión o no ponen en marcha ninguna estrategia compensatoria (Alonso Tapia y Cols., 1992; Solè, 1992; Repetto, Téllez y Beltrán, 2002). • No elaboran la información y, por lo tanto, no se integra en sus conocimientos ni se generaliza a otras situaciones (Palinskar y Brown, 1984; Solè, 1992). • No evalúan ni la eficacia ni la efectividad de su actividad lectora (Solè, 1992; Solè y Teberosky, 2001; González y Romero, 2001). • No son conscientes de los objetivos, procesos o estrategias implicadas en la lectura (Alonso Tapia y Carriedo, 1996; Repetto, Téllez y Beltrán, 2002; González y Romero, 2001). • Deficiencias en la progresión temática de los textos (Alonso Tapia y Carriedo, 1996). • Actitud negativa hacia la lectura (Romero, 1990; Mora, 1991; Sánchez, 1999; Mora Merchán y Mora, 2000). • Uso de estrategias de listado (reproducción de ideas inconexas y no captación del sentido global del texto) (Scardamalia y Bereiter, 1984; González y Romero, 2001). • Uso de estrategias de "suprimir y copiar" (eliminar la información redundante y copiar el resto de forma literal) (Brown, Day y Jones, 1983; Winograd, 1984). • Mala autorregulación del proceso lector (Sánchez, 1998; 1999; Alonso Tapia y Cols., 1992; González y Romero, 2001).

simultánea actividades a su vez también complejas y que se ven altamente influenciadas —cuando no condicionadas— por factores que escaparían a un primer análisis basado primordialmente en aspectos excluyentes, ya sean estos cognitivos, emocionales o de cualquier otra índole. A modo de ejemplo práctico, bastaría pensar unos instantes en las habilidades y subhabilidades que en estos precisos momentos estamos poniendo en práctica cuando leemos este texto, y cómo están influyendo otros elementos de tipo emocional o social en nuestra lectura. Consecuente con la afirmación anterior, y como en muchos otros procesos, el de la lectura exige también para su análisis exhaustivo, un recorrido minucioso a través de los elementos que los seres humanos ponemos en juego de forma simultánea cuando leemos un texto. Este análisis de la lectura ha tenido en nuestro caso el objetivo de diseñar un dispositivo de evaluación que fuera sensible a cada uno de estos procesos analizados que intervienen en la comprensión.

Podríamos decir que, en principio, las acciones ejecutadas durante el proceso lector van encaminadas a la consecución de varios objetivos: por un lado, algunas están implicadas en la asignación de un significado a los símbolos arbitrarios que históricamente se han construido, poniendo en marcha, entre otros, procesos de identificación, asociación, discriminación visual y fonológica, etc.; otras acciones están más implicadas en la conjugación de las diversas segmentaciones de las que consta un texto, para su comprensión e interpretación global; hay un tercer grupo de acciones encaminadas a regular y controlar el propio proceso de la lectura; y, por último, están aquellas que se dirigen a la conexión del contenido del texto con las experiencias y los conocimientos previos.

Estas acciones, como venimos comentando, tienen asimismo, la doble característica de ponerse en práctica de forma *interactiva* (Adams, 1980; Rumelhart, 1980; Miller, 1985; Alonso y Mateos, 1985; Sánchez, 1990; Alonso Tapia y Cols. 1992; Solè, 1987, 1992; Maruny, Ministrál y Miralles, 2001) e *inmediata* (Sánchez, 1990). Quizás el concepto más relevante y que sintetiza más adecuadamente el tipo de relaciones que existen entre los distintos procesos y elementos puestos en juego en el hecho subjetivo de la comprensión de la lectura, sea el de *interactividad*. Con este concepto se hace referencia a que los procesos que se recogen en la figura 1 no tienen lugar de forma lineal o sucesiva; por el contrario su actuación se da de forma simultánea. Como se observa en la figura, todos influyen en la construcción de un *modelo mental de comprensión* (Valiña, Bernal y de Vega (1990), es decir, en el nivel más alto de comprensión, que representa aquello que el sujeto ha aprendido del texto, donde se extrae e interpreta la idea esencial o «teoría» actual del lector acerca de lo leído. Los otros niveles de comprensión descritos: *micronivel de comprensión* o texto base, construido con las proposiciones de un texto, donde se accede al conocimiento de lo que propone el texto, sin proceder a su análisis exhaustivo; y *macronivel de comprensión*, que consiste en el resumen de la estructura proposicional, donde se construyen conclusiones y se crean los esquemas mentales acerca del contenido y la estructura textual, comparten la característica básica de interactividad de los procesos puestos en juego, pero, obviamente, no la participación de todos los procesos descritos. A su vez, todos los procesos que influyen en la construcción de la comprensión global de un texto se ven influidos por los demás, conformando una red interactiva y sistémica de relaciones que explica la comprensión. Ahora bien, ¿qué papel juegan cada uno de los distintos procesos y elementos en la explicación del acceso a la lectura? ¿Son imprescindibles todos? ¿O se puede llegar a la comprensión de textos sin el concurso de alguno? Tratando de responder a estas preguntas, diferentes estudios de análisis componencial

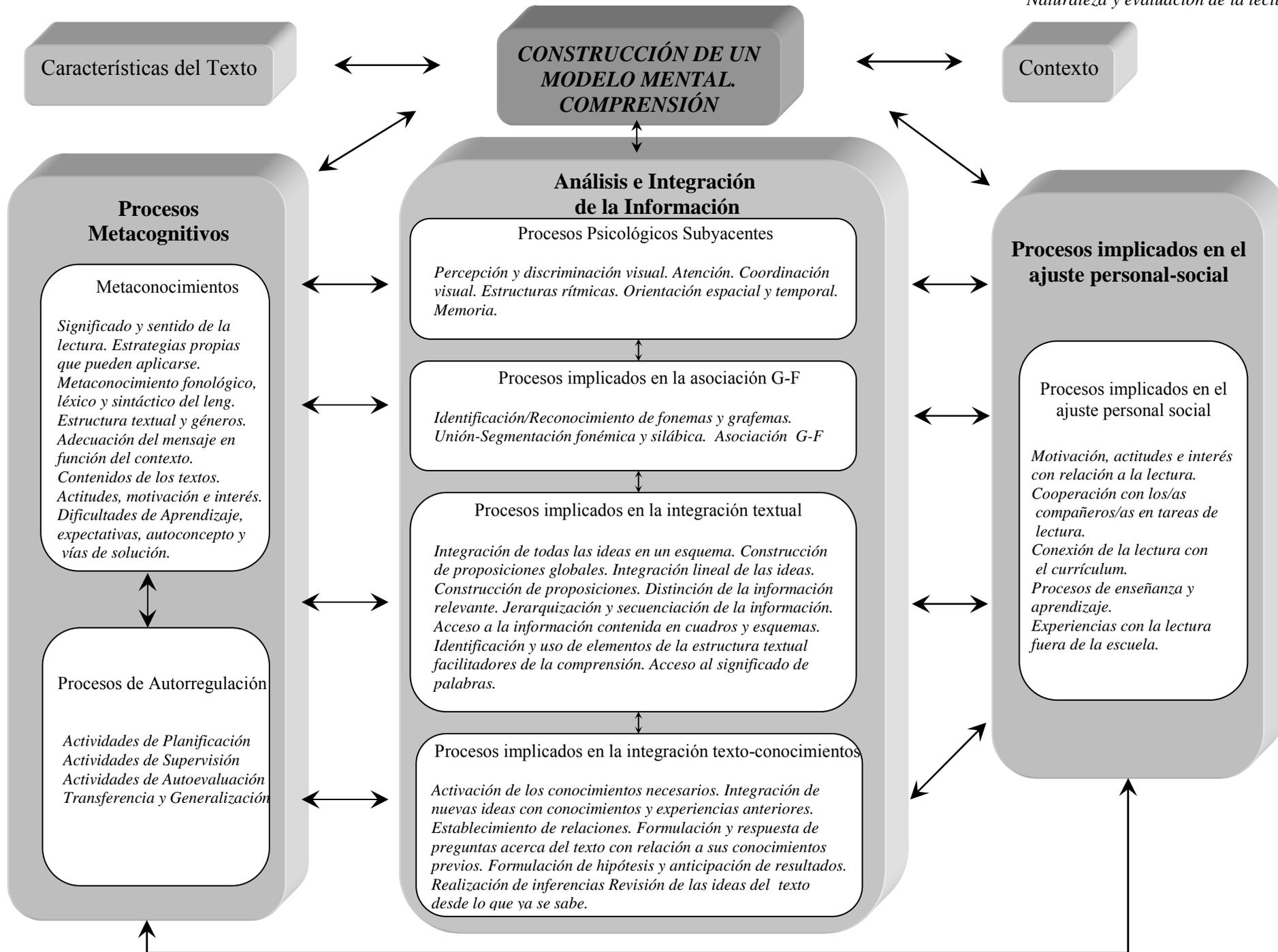


Figura 1. Relaciones entre los procesos implicados en la lectura

de la habilidad lectora pusieron de manifiesto que los componentes principales del proceso lector serían los procesos de decodificación y los de comprensión verbal (Jackson y McClelland, 1979; Palmer, McLeod, Hunt y Davidson, 1985; Stanovich, Cunningham y Cramer, 1984). Posteriormente se ha revelado que, tanto en unos procesos como en otros, juega un papel especialmente relevante la puesta en práctica de los procesos metacognitivos: los metaconocimientos y las habilidades de autorregulación (Mora, 1991; Vidal-Abarca y Gilbeert, 1991; Alonso Tapia y Cols., 1992; Alonso Tapia y Carriedo, 1996; Jiménez y Ortiz, 1998; Sánchez, 1999; Moreno y García, 2000).

Existen, por otra parte, algunas matizaciones a las afirmaciones anteriores relativas a la interacción de elementos en el proceso. Al parecer, y aunque hemos apuntado que al analizar la lectura y los procesos y elementos que la componen, analizamos procesos y elementos que se relacionan actuando de forma interactiva, el acceso a la comprensión de textos requiere necesariamente de la información procedente de los procesos denominados *ascendentes*, es decir, aquellos que en último término buscan descodificar unos símbolos escritos para asignarles un significado. Esto no contradice lo anterior, pero sí limita de alguna forma la interacción existente entre estos procesos, ya que cualquier problema o inadecuada ejecución en los procesos implicados en la asociación de símbolos lingüísticos con un significado, conllevaría una afectación del resto de procesos y elementos implicados (Just y Carpenter, 1987; Alegría, 2006); bajo nuestro punto de vista, no tanto en el sentido que se vean afectados en sí, sino más bien en que no tienen la oportunidad de llevarse a cabo. Por ejemplo, hay alumnos que manifiestan graves problemas en los procesos de asociación grafemas-fonemas y sin embargo, una vez que, con grandes dificultades, han descodificado una palabra, el nivel de comprensión en cuanto al acceso al significado de la palabra es bueno. Relacionado con lo anterior, parece ser que en los estadios iniciales del aprendizaje de la lectura, el desarrollo de procesos implicados en la asociación de grafemas y fonemas, y, particularmente, la utilización de la llamada ruta fonológica de acceso al significado, es particularmente relevante (Henry, 1993; Perfetti y Hogaboam, 1975; Domínguez y Cuetos, 1992; Jiménez y Rodrigo, 1994).

A pesar de las afirmaciones expuestas y de que es cierto, por otra parte, que la existencia de problemas en determinados procesos que contribuyen a optimizar el proceso y los resultados de la lectura, no conllevan una afectación general del proceso lector —tales son algunos problemas leves en procesos de tipo viso-perceptivo o visomotor (González, 1984) o también algunos procesos de ajuste personal-social—; si entendemos la lectura como un proceso primordialmente encaminado a la comprensión, y que se desarrolla y sustenta en una estructura sistémica de relaciones entre los subprocesos y elementos implicados, la presencia de problemas en cualquiera de los procesos que se analizarán en este trabajo y que quedan recogidos en la figura 1, afectará al proceso completo y, por el contrario, la ejecución adecuada de todos estos procesos coexistiría en una lectura hábil y eficaz.

Aunque, como muestran numerosos trabajos (Ferreiro y Teberosky, 1982; Ferreiro, 1997; Solè y Teberosky, 2001; Maruny, Mistral y Miralles, 2001) relacionados con la enseñanza y el aprendizaje del lenguaje escrito en niños pre-lectores y en lectores recién iniciados, la utilización de procesos de construcción e integración de significados, integración del texto con los conocimientos y las experiencias previas —incluyendo la elaboración de hipótesis e inferencias, así como la propia reflexión y manipulación de

segmentos estructurales de los mensajes escritos—, tiene lugar desde los primeros momentos del aprendizaje de la lectura; a medida que la habilidad lectora se desarrolla, aumentaría el uso de estos procesos denominados de orden *descendente*. Con la práctica, los procesos de asociación grafema-fonema, a los que antes hacíamos alusión, y, en definitiva, aquellos encaminados al reconocimiento de palabras, se irían automatizando, pudiendo dedicarse los recursos atencionales a los procesos comprensivos de más alto nivel. Estos procesos de comprensión, como ya hemos mencionado, se realizan conjugando e integrando diversos elementos textuales y contextuales. La interpretación del lenguaje escrito tiene lugar transformando los símbolos lingüísticos en una representación mental aún más abstracta si cabe. Estaríamos pasando del lenguaje al pensamiento. Esta comprensión puede ser de información explícita (contenida en el texto) o implícita (necesita ser inferida a raíz de la información textual). Intervienen por tanto multitud de factores, desde habilidades para asociar grafemas y fonemas pasando por aquellas que permiten una estructuración adecuada del mensaje escrito, hasta los conocimientos que tiene el lector sobre el texto, desde las habilidades para realizar inferencias hasta la riqueza y amplitud del vocabulario, desde los procesos de control y autorregulación de la propia comprensión hasta la capacidad de elaborar hipótesis y plantearse dudas y preguntas acerca del contenido del texto, etc. Una evaluación que pretenda obtener información útil de cara a una acción educativa eficaz y orientada a la solución de las dificultades encontradas, debe contar, por una parte, con la *presencia de estos procesos* en la lectura y, por consiguiente, en el dispositivo de evaluación; con la característica de *interactividad* que presentan tales procesos; con una manera *sistémica* de relacionarse; y, por último, con la relevancia de evaluar precisamente el *proceso de leer*.

1.2 La evaluación de la lectura y la comprensión de textos

Introducción

La enseñanza y el aprendizaje de la lectura, entendida ésta en los términos descritos anteriormente de proceso interactivo y dinámico, que es construido socialmente en contextos de interacción y que consiste básicamente en la construcción y representación global de significados, requiere un proceso de evaluación interesado en determinar las necesidades del alumnado y las posibles dificultades encontradas en el acceso a la comprensión para poder dar solución a las mismas. Esta evaluación, necesaria para poder afrontar la enseñanza de la lectura y, sobre todo, de las estrategias de lectura, debe conocer además la naturaleza de aquello que se pretende enseñar; en este caso, debe conocer en primer lugar que la lectura como tal es un *proceso*, es decir, una actividad que desarrollamos en el tiempo y que, por lo general, va conformando su propio sentido a medida que avanzamos en la actividad. Esta naturaleza procedimental de la lectura — que, por otra parte, no es la única, ya que como exponíamos al recoger los puntos de la concepción de la lectura que serviría de marco a las propuestas recogidas en este trabajo, ésta puede ser susceptible de ser llevada a la conciencia y de ser objeto de reflexión no sólo como proceso, sino también como objeto de conocimiento— se refleja igualmente cuando se tratan de analizar los procesos y elementos que forman parte de ella. De este modo, la lógica y la naturaleza del objeto que tratamos nos llevaría a la necesidad de evaluar los distintos procesos implicados en la lectura: lingüísticos, cognitivos, metacognitivos, socio-emocionales y contextuales; y de evaluar, así mismo, el proceso mismo de leer. Esta evaluación de procesos nos lleva, a su vez, a considerar el conjunto de estrategias y conocimientos que se ponen en juego en la construcción de

una representación de significados a raíz de la lectura de un texto. En síntesis, la enseñanza de la lectura llevaría a la necesidad de la evaluación, ésta nos conduciría a la necesidad de evaluar procesos, y esto a la necesidad de tener en cuenta las estrategias que el lector debe poner en marcha para acceder a la comprensión.

Pero vayamos más despacio; aunque podemos encontrar diversas acotaciones del término, podríamos definir básicamente la *evaluación educativa* como un elemento integrante del proceso de enseñanza y aprendizaje a través del cual obtenemos información acerca del grado de consecución de determinados objetivos educativos, con el doble fin de valorar el aprendizaje efectuado, por una parte, y de que sirva para la toma de decisiones encaminada a regular la acción educativa de forma ajustada a las necesidades observadas, por otra (Coll, Martín y Onrubia, 2001; Ontoria, 1993; Coll, Barberà y Onrubia, 2000; Coll y Onrubia, 1999).

Se han distinguido clásicamente tres tipos de evaluación escolar de los aprendizajes, ligadas habitualmente a determinadas funciones y también a decisiones pedagógicas. En primera instancia la *evaluación inicial o diagnóstica* tiene lugar al comienzo del proceso y suele estar encaminada a establecer una valoración inicial sobre las capacidades del alumnado, bien para ajustar la enseñanza a las necesidades detectadas, o bien para orientar diferencialmente a éste hacia objetivos y contenidos curriculares diversos en función de las capacidades diagnosticadas. Lógicamente, estos dos caminos que se abren a raíz de una evaluación inicial suelen ser indicadores, a su vez, de sendas formas diferentes, y a menudo antagónicas, de entender la evaluación y los procesos mismos de enseñanza y aprendizaje. Por otra parte, la *evaluación continua*, que suele tener asociado un carácter formativo o regulador de la enseñanza y también formador (Nunziati, 1990), hace hincapié en la recogida de información continua durante el desarrollo mismo de las actividades de enseñanza y aprendizaje; este modelo suele conectar, de este modo, las actividades encaminadas a la consecución de los objetivos de aprendizaje con las actividades encaminadas a la evaluación que del proceso mismo se efectúa. Se prima la evaluación del proceso educativo y suele estar en la base de la ayuda a la toma de decisiones acerca de la mejora del proceso. Se relaciona con la posibilidad de regulación continua y ajustada de los procesos de enseñanza a las necesidades y motivación del alumnado, así como con la posibilidad de introducir elementos de autorregulación de los aprendizajes por parte del propio alumnado (Mauri y Mirás, 1996). La evaluación continua, por tanto, se desarrolla mediante una evaluación de procesos, asumiendo necesariamente la naturaleza dinámica de estos procesos. En tercer lugar, la *evaluación sumativa o final* se identifica con la valoración que se realiza al final de una actividad o conjunto de actividades de enseñanza y aprendizaje acerca de la consecución de determinados objetivos de aprendizaje (Coll, Martín y Onrubia, 2001; Mirás y Solè, 1990). Suele estar asociada a una función acreditativa o diagnóstica de los aprendizajes, por lo que puede estar en la base de la orientación diferencial del alumnado hacia vías curriculares distintas, al igual que ocurría con la evaluación inicial; de otro modo, puede estar también al servicio de una concepción menos estática y más reguladora del proceso educativo, cuando se utiliza de forma asidua y durante el desarrollo del proceso como fuente de información para ajustar la enseñanza a las necesidades educativas del alumnado.

Enfoques evaluativos: El enfoque del test y la evaluación integrada en el proceso de enseñanza y aprendizaje

Al hilo de esta distinción entre los distintos tipos de evaluación educativa, podemos deducir, con la ayuda de Coll, Barberà y Onrubia (2000) dos grandes enfoques o *culturas evaluativas* en función de (a) las prácticas de evaluación de los docentes, (b) las decisiones adoptadas en función de los resultados de estas prácticas, así como de (c) las funciones que las mismas tienen dentro —o fuera— del proceso educativo. En primer lugar, destacan estos autores una *cultura del test* caracterizada básicamente por otorgar a la evaluación una función esencialmente social, desligando las tareas de enseñanza y aprendizaje de aquellas que valorarán el grado de aprendizaje alcanzado. La función social, de certificación o acreditación, de estas prácticas evaluativas se fundamenta, así mismo, en una concepción cuantitativa y acumulativa del aprendizaje, que engarza con la tradición asociacionista y los modelos conductistas en psicología. Se asume que el resultado del aprendizaje se debe, por una parte, a la reproducción del conocimiento y de las habilidades por parte del alumnado, secuenciados lineal y jerárquicamente por el docente. En este sentido es posible hablar de «niveles de aprendizaje» alcanzados y se situaría el énfasis en la medición cuantitativa de los logros. Así mismo, en el marco general de esta *cultura del test*, las diferencias individuales y la atención educativa a estas diferencias suelen ser concebidas con un carácter estático, en el sentido de que las diversas características individuales como la inteligencia, la motivación o la capacidad de aprendizaje, serían elementos o rasgos prefijados e inmutables y, por lo tanto, la evaluación escolar estaría al servicio de una función selectiva de la educación, aceptando como inevitable el que un determinado número de alumnos y alumnas fracasasen debido a su falta de capacidad, inteligencia o motivación.

Por otra parte, la segunda de estas culturas evaluativas vincula las prácticas de evaluación con las actividades de enseñanza y aprendizaje, centrándose de este modo más en una función pedagógica que social. Se entiende la evaluación como parte integrante del proceso educativo y, por tanto, las actividades encaminadas a la obtención de información sobre el aprendizaje y, en este caso, también sobre la propia enseñanza, tienen lugar a lo largo del desarrollo del proceso, por lo que se relacionaría este segundo enfoque con la evaluación continua o formativa, anteriormente descrita. Se trata de obtener información útil y relevante —sobre todo de tipo cualitativo— sobre las competencias del alumnado con relación a los aprendizajes realizados, con vistas al ajuste de la acción educativa que permita la mejora de todo el proceso. Se da prioridad en las tareas de evaluación propuestas desde este enfoque a aquellas que destacan un *proceso de aprendizaje* y la capacidad del alumno o alumna para poner en práctica habilidades de pensamiento como comprender, inferir, elaborar, redactar, planificar, supervisar, etc. Esta cultura de la evaluación suele relacionarse con posiciones teóricas enraizadas en la psicología cognitiva y el constructivismo. En este caso, el aprendizaje se considera un proceso activo de construcción compartida de conocimientos, procedimientos y actitudes con base en representaciones previamente construidas y con la ayuda de otros más expertos. Desde este enfoque, la evaluación aportaría información sobre el proceso de construcción de estas representaciones para orientar y ajustar la ayuda pedagógica a la re-construcción de las mismas en una continua optimización del proceso.

Así mismo, las diferencias individuales y la atención a la diversidad son concebidas en una forma significativamente distinta a la anterior; en primer lugar, las características

individuales a las que aludíamos: la inteligencia, la motivación o la capacidad de aprendizaje, son entendidas como elementos susceptibles de ser modificados con base en la experiencia y en el aprendizaje escolar, acentuando el carácter *dinámico* de estas características individuales; en segundo lugar, desde este posicionamiento se está asumiendo explícita e implícitamente una concepción *interaccionista* de las diferencias. En efecto, una evaluación continua, que se centra en el proceso mismo de enseñanza y aprendizaje, necesariamente constatará que los resultados y el desarrollo de las actividades propuestas tiene mucho que ver con la calidad de los recursos de los que se disponga, así como de las propuestas educativas que se realizan, concluyendo más fácilmente que las diferencias individuales están determinadas por la interacción de factores inherentes al individuo y factores relativos al proceso mismo de enseñanza y aprendizaje. Al mismo tiempo, esta concepción interaccionista de las diferencias lleva en buena lógica a una concepción a su vez dinámica de las características individuales, en el sentido de que éstas son objeto *continuo* de una acción educativa encaminada a facilitar la construcción y reconstrucción de representaciones y significados elaborados, lo cual no tendría mucho sentido si se creyera que las características aludidas son inamovibles e inmodificables. En líneas generales, dentro de estos dos posicionamientos teóricos sobre la evaluación, podríamos encuadrar las distintas propuestas de evaluación de la comprensión lectora. A continuación, nos adentraremos con mayor detalle en las pruebas y procedimientos desarrollados desde cada uno de estos enfoques.

La evaluación de la lectura y la comprensión desde el enfoque del test

La denominada *cultura del test* no solo es representativa de la gran mayoría de las prácticas evaluativas en los centros escolares, sino que también es inspiradora de gran parte de los procedimientos para evaluar la comprensión que se han ido gestado en las últimas décadas a raíz de la relevancia, cada vez mayor, otorgada a esta área dentro del proceso educativo escolar. Diversos autores son los que han acometido la misión de realizar revisiones de las pruebas de evaluación de la comprensión para nuestro contexto, basadas en esta cultura del test (Donoso, 1994; Jiménez, Rodrigo, Ortiz y Guzmán, 1999; Defior y Ortuzar, 1993; Moreno y García, 2000; Repetto, Téllez y Beltrán, 2002). Estas pruebas se caracterizan, entre otros aspectos, por el énfasis en la medición cuantitativa de los resultados, la búsqueda de objetividad y de fiabilidad estadística en las respuestas obtenidas, la fundamentación teórica orientada a la comprobación o la confirmación de hipótesis o diagnósticos previamente fijados, la orientación hacia el resultado o la asunción de una naturaleza estática de las aptitudes individuales y de la realidad (Cook y Reichardt, 1986). Respondiendo en mayor o menor medida a estas características, exponemos a continuación en el cuadro 2 los puntos esenciales de algunas de estas pruebas, siguiendo en parte los trabajos de revisión citados anteriormente. El encuadre en las tres grandes categorías propuestas por Donoso (1994) que adoptamos y adaptamos para clasificar estas pruebas, no siempre puede ser del todo ajustado, pero puede ser útil a los propósitos de sintetizar la información y facilitar la comprensión al lector.

Cuadro 2. Pruebas estandarizadas de evaluación de la lectura

Pruebas de evaluación de prerrequisitos o habilidades consideradas necesarias para el aprendizaje de la lectura			
<i>Evaluación de las habilidades metalingüísticas</i>			
Nombre	Autores y edición	Dimensiones evaluadas	Aplicación
Prueba de Conocimiento sobre el Lenguaje Escrito (CLE)	M ^a R. Ortiz y J.E. Jiménez (1993) Madrid. TEA ediciones.	*Conciencia sobre las funciones y características del lenguaje escrito	Educación Infantil
Prueba de Segmentación Lingüística (PSL)	J.E. Jiménez y M ^a R. Ortiz (1998). <i>Conciencia fonológica y aprendizaje de la lectura</i> . Madrid. Síntesis.	*Conciencia fonológica. Se evalúan la conciencia fonémica, silábica y léxica. Actividades de segmentación, omisión, síntesis, reconocimiento y aislamiento de unidades fonológicas.	2º curso de E. Infantil
Evaluación de la Conciencia Intrasilábica (PCI)	J.E. Jiménez y M ^a R. Ortiz (1998). <i>Conciencia fonológica y aprendizaje de la lectura</i> . Madrid. Síntesis.	*Conciencia intrasilábica. Se evalúa la capacidad de reconocimiento de unidades intrasilábicas (onset y rima).	2º curso de E. Infantil
Prueba de Conciencia Fonémica (PCF)	J.E. Jiménez y M ^a R. Ortiz (1998). <i>Conciencia fonológica y aprendizaje de la lectura</i> . Madrid. Síntesis.	*Conciencia fonémica. Actividades de síntesis, aislamiento, segmentación y omisión de fonemas.	1º ciclo de E. Primaria y alumnado mayor con N.E.E.
Test de Habilidades Metalingüísticas (THM)	P.F. Gómez, J. Valero, R. Buades y A. Pérez (1995). Madrid. EOS	*Conciencia fonológica. Actividades de Identificación-Reconocimiento fonémico, Discriminación-Sustitución fonémica y Segmentación silábico-fonémica.	2º ciclo de E. Infantil
<i>Evaluación general de habilidades y prerrequisitos para la madurez lectora</i>			
Nombre	Autor y edición	Dimensiones evaluadas	Aplicación
Test A.B.C.	L. Filho (1973) <i>Test A.B.C.</i> Buenos Aires. Kapelusz	*Coordinación viso-motora *Resistencia a la inversión en la copia de figuras *Memoria visual y auditiva *Coordinación audiomotora *Capacidad de pronunciación *Vocabulario *Comprensión	E. Primaria

Batería predictiva del aprendizaje de la lectura INIZAN	A. Inizan (1983) <i>Cuándo enseñar a leer</i> . Madrid. Aprendizaje Visor.	*Aspectos predictivos *Lectura de palabras familiares y no familiares *Comprensión de lectura silenciosa	Entre 5 y 7 años
Aptitudes en E. Infantil (AEI)	M.V. de la Cruz (2000) Madrid. TEA ediciones.	*Aptitudes verbales *Aptitudes numéricas *Aptitudes espaciales *Memoria *visomotricidad	2º ciclo de E. Infantil
Batería evaluativa de las habilidades necesarias para el aprendizaje de la lectura y la escritura (BEHNALE)	J.A. Mora (1993) Madrid. TEA ediciones.	*Articulación *Percepción y discriminación auditiva *Estructuración espacio-temporal *Memoria visual inmediata *Coordinación visomotora *Memoria motora *Percepción y discriminación visual *Vocabulario *Memoria auditiva y lógica inmediata *Discriminación lateral	Entre 5 y 6 años
Pruebas de evaluación de estados iniciales de la lectura			
Nombre	Autor y edición	Dimensiones evaluadas	Aplicación
Test de Análisis de la Lecto-Escritura (TALE)	J. Toro y M. Cervera (1984) <i>Test de Análisis de lecto-escritura</i> . Madrid. Aprendizaje Visor.	Determinación de niveles y características específicas de la lectura y la escritura. *Comprensión lectora *Exactitud lectora *Escritura	1º a 4º de E. Primaria
Exploración de las dificultades individuales de lectura (EDIL 1)	M.D. González Portal (1985) Madrid. TEA ediciones.	*Exactitud lectora *Comprensión lectora *Velocidad lectora *Nivel global de lectura	Entre 5 y 7 años
Pruebas de lectura, niveles 1 y 2 (LECTURA)	M.V. de la Cruz (1988) Madrid. TEA ediciones.	*Aptitud lectora *Conocimiento de lectura	Nivel (1) Entre 5 y 6 años Nivel (2) 7 años
Evaluación de la Comprensión Lectora (ECL 1)	M.V. de la Cruz (1995) Madrid. TEA ediciones.	*Capacidad para captar el sentido de textos de uso habitual *Analizar aspectos sencillos típicos de diferentes tipos de textos *Conocimientos del significado de palabras, sinónimos y antónimos *Comprensión del significado de frases *Capacidad de integración de información textual	ECL (1) Entre 6 y 9 años (1º-4º de E.P.)

Batería de evaluación de los procesos lectores de los niños de E.P. (PROLEC)	F. Cuetos, B. Rodríguez y E. Ruano (1996). <i>Evaluación de los procesos lectores (PROLEC)</i> . Madrid. TEA ediciones.	*Identificación de letras *Reconocimiento de palabras *Procesos sintácticos *Procesos semánticos *Capacidad global lectora	1º-4º de E. Primaria
Pruebas de evaluación de estadios o niveles superiores de comprensión			
Nombre	Autor y edición	Dimensiones evaluadas	Aplicación
Evaluación de la Comprensión Lectora (ECL 2)	M.V. de la Cruz (1995) Madrid. TEA ediciones.	*Capacidad para captar el sentido de textos de uso habitual *Analizar aspectos sencillos típicos de diferentes tipos de textos *Conocimientos del significado de palabras, sinónimos y antónimos *Comprensión del significado de frases *Capacidad de integración de información textual	ECL (2) Entre 8 y 11 años (3º-6º de E.P.)
Prueba de Comprensión lectora	A. Lázaro Martínez (1980). Madrid. TEA ediciones.	*Forma de expresión *Ritmo y forma de percepción *Nivel y tipo de comprensión	A partir de 8 años
Prueba de Comprensión Verbal (TCV) Nivel 2 y 3	E. Díez Fernández (1980). <i>Test de comprensión verbal</i> . Madrid. Publicaciones ICCE.	*Lectura comprensiva *Vocabulario *Reconocimiento del significado de las palabras	TCV (2) Entre 9 y 11 años TCV (3) Entre 12 y 13 años
Pruebas de Comprensión Lectora (CLT-CLOZE)	A. Suárez y P. Meara (1985). <i>Dos pruebas de comprensión lectora. Procedimiento CLOZE</i> . Madrid. TEA ediciones.	Comprensión lectora	Adolescentes y adultos
Evaluación de los procesos lectores en alumnos de 3º ciclo de E.P. y E.S.O. (PROLEC-SE)	J.L. Ramos y F. Cuetos (2000). <i>Evaluación de los procesos lectores (PROLEC-SE)</i> . Madrid. TEA ediciones.	*Reconocimiento y lectura de palabras y pseudopalabras *Signos de puntuación *Procesos sintácticos y semánticos *Estructura textual *Capacidad lectora global	3º ciclo de E.P. y E.S.O.
Batería de Supervisión y Regulación de la Comprensión Lectora SURCO	J.A. Tapia, N. Carriedo y M.M. Mateos (1992). <i>Leer, comprender y pensar: nuevas estrategias y técnicas de evaluación</i> . Madrid. CIDE.	*Autorregulación de la comprensión lectora *Detección de fallos de comprensión *Conocimiento de estrategias reguladoras a aplicar en situaciones específicas *Comprensión basada en el uso de estrategias	6º de E.P. y 1º y 2º de E.S.O.

Evaluación de la capacidad de comprender y resumir lo importante. Bateria IDEPA	J.A. Tapia, N. Carriedo y E. González (1992). <i>Leer, comprender y pensar: nuevas estrategias y técnicas de evaluación</i> . Madrid. CIDE.	*Vocabulario *Detección de ideas principales - Representación - Conocimiento estratégico - Identificación de ideas principales *Elaboración de resúmenes - Representación - Estrategias de resumen - Identificación del mejor resumen	6º de E.P. y E.S.O.
Evaluación de la comprensión y el pensamiento crítico. Bateria CRITEX	J.A. Tapia y F. Gutiérrez Martínez (1992). <i>Leer, comprender y pensar: nuevas estrategias y técnicas de evaluación</i> . Madrid. CIDE.	*Lectura crítica *Razonamiento básico *Razonamiento silogístico *Razonamiento condicional *Razonamiento sobre textos naturales	6º de E.P. y E.S.O.
Evaluación de la Comprensión Lectora. Pruebas ACL	G. Català, M. Català, E. Molina y R. Monclús (2001). <i>Evaluación de la comprensión lectora. Pruebas ACL</i> . Barcelona. Graò.	Comprensión lectora *Comprensión literal *Reorganización de la información *Comprensión inferencial *Comprensión crítica *Intencionalidad del autor	ACL1 1º de E.P. ACL2 2º de E.P. ACL3 3º de E.P. ACL4 4º de E.P. ACL5 5º de E.P. ACL6 6º de E.P.

Lógicamente no todas estas pruebas parten de una concepción idéntica de la lectura; es más, entre ellas existen notables diferencias. Algunas de ellas parten de posicionamientos teóricos interactivos en cuanto a la explicación de la lectura y de sus dificultades de aprendizaje, e incluso desde algunas de estas propuestas se aboga por una evaluación de procesos que complemente la información recogida con la pasación de los tests estandarizados. Como tendremos oportunidad de analizar posteriormente de forma más detallada, algunas pruebas de las recogidas en la revisión se centran en la valoración de estrategias metacognitivas de autorregulación del proceso de comprensión, aunque desde un enfoque cuantitativo de respuestas cerradas. Sin embargo, es interesante comprobar cómo se han ido modificando y enriqueciendo algunos de estos procedimientos en la dirección de una evaluación que tenga más en cuenta los distintos procesos que la investigación ha ido desvelando como componentes del proceso de comprensión. Pruebas de evaluación de la comprensión como PROLEC, en sus dos versiones; las baterías SURCO, IDEPA y CRITEX, las pruebas CLT-CLOZE o las pruebas ACL, son ejemplos en los que, aunque prima un enfoque cuantitativo y valorativo de resultados, se incorpora, al menos en sus propuestas teóricas, la necesidad o la conveniencia de analizar cualitativamente los resultados con el fin de obtener información que pueda ser más valiosa para la intervención.

La evaluación de la lectura y la comprensión desde el enfoque integrador

Por otra parte, se encuentran aquellos procedimientos de evaluación que se han ido desarrollando bajo el paraguas de un paradigma de corte más cualitativo y que se relacionaría con el segundo tipo de cultura evaluativa al que hacíamos referencia (Coll,

Barberà y Onrubia, 2000). Dentro de este grupo podemos comenzar citando todas aquellas actividades integradas intencionalmente por los docentes en el proceso de enseñanza y aprendizaje, de las cuales se extrae información contextualizada acerca de la marcha del proceso y del grado en que se están cumpliendo los objetivos propuestos. Algunas de estas actividades, como realizar una exposición de un trabajo, componer un texto a raíz de un centro de interés trabajado, elaborar una síntesis de las ideas fundamentales tras una visita organizada, participar en una puesta en común tras haber debido preparar un tema determinado, etc.; sirven al propósito de evaluar de forma continua y contextualizada la competencia del alumnado, centrando el interés tanto en el proceso como en el producto y con una clara concepción dinámica de los procesos educativos. Desde una concepción constructivista del proceso de enseñanza y aprendizaje escolar se entiende así la evaluación de la lectura (Solè, 1992; Coll, Barberà y Onrubia, 2000; Maruny, Ministrál y Miralles, 2001).

Algunas de las propuestas efectuadas desde este modelo destacan la importancia de aprovechar textos y contextos habituales para valorar el grado de dominio de las estrategias necesarias para comprender (Bofarull, 2001). También se hace hincapié en la posibilidad de utilizar cualquier área de conocimiento para evaluar la lectura y no necesariamente el área de Lengua. El contexto es sumamente importante desde esta concepción de la educación escolar; Solè (1992), entre otros, habla de contextos funcionales y significativos para el alumnado a la hora de la evaluación. Se aboga por la utilización de textos completos, ya que no se trata de valorar una habilidad específica de forma aislada, sino la capacidad de construir una representación global del significado de un texto, debiendo para ello conjugar distintas estrategias de lectura. Se entiende claramente la evaluación como un proceso continuo orientado a la regulación y mejora del aprendizaje y de la enseñanza y, por tanto, centrada fundamentalmente en el proceso. Colomer y Camps, citadas en Solè (1992), realizan una propuesta sobre aquellos elementos acerca de los que sería necesario obtener información en un proceso de evaluación de la lectura. Estos son: (a) la actitud emocional con que el lector se enfrenta al texto, (b) el grado en que la lectura realizada se adecua a los objetivos propuestos, (c) el grado en que el alumno puede manejar las fuentes escritas, (d) el proceso de construcción del significado, (e) el grado en que el lector controla su propio proceso de comprensión y (f) sobre la lectura oral y la velocidad lectora.

Por su parte, Bofarull (2001), recogiendo las ideas anteriores, propone, desde esta concepción constructivista del proceso de enseñanza y aprendizaje de la lectura, la *observación* como instrumento preferente y valioso para obtener información sobre la competencia lectora del alumnado. Esta observación debe ser realizada, en línea con lo que señalábamos anteriormente, sobre actividades de lectura contextualizadas, y que sean significativas para el alumnado. Esta autora propone una pauta de observación para evaluar la lectura, que recogemos a continuación en el cuadro 3. La propuesta desarrolla básicamente los puntos a tener en cuenta que eran listados con anterioridad. Recogiendo varias cuestiones a modo de indicadores de la puesta en práctica de las distintas estrategias o procesos componentes de la lectura, distribuye éstos en tres momentos: antes de la lectura, durante la lectura y después de la lectura; y aporta unos comentarios sobre la pauta de observación, que desarrollan cada uno de los puntos contenidos en la misma y facilita su aplicación por parte del profesorado.

Cuadro 3. Pauta de observación de la comprensión lectora.
(Tomado de Bofarull, 2001).

PAUTA DE OBSERVACIÓN DE LA COMPRENSIÓN LECTORA	
ANTES DE LA LECTURA	OBSERVACIONES
<ul style="list-style-type: none"> • ¿Qué actitud emocional presenta el lector o lectora? • ¿Cuál es el objetivo de la lectura? • ¿Activa el conocimiento previo? • ¿El lector o lectora realiza hipótesis y predicciones sobre el contenido del texto? 	
DURANTE LA LECTURA	OBSERVACIONES
<ul style="list-style-type: none"> • ¿Usa las señales del texto para construir significados? • ¿Verifica las hipótesis, plantea otras nuevas y, si existe algún error, reflexiona para encontrar su causa? • Ante las dificultades, ¿Qué recursos utiliza para superarlas? 	
DESPUÉS DE LA LECTURA	OBSERVACIONES
<ul style="list-style-type: none"> • ¿Identifica el tema? • ¿Identifica la idea principal? • ¿Es capaz de hacer una comprensión literal? • ¿Es capaz de hacer una comprensión inferencial o interpretativa? • ¿Es capaz de hacer una comprensión profunda o crítica? • ¿Es capaz de hacer un resumen coherente? • ¿La velocidad lectora es adecuada? 	

Otras autoras (Maruny, Ministrál y Miralles, 2001), también desde el contexto escolar y en una serie de volúmenes titulados genéricamente *Escribir y Leer*, realizan otra propuesta de evaluación que prima la contextualización de los instrumentos de evaluación, la observación en contextos naturales y significativos para el alumnado y el registro de los conocimientos y el uso de lenguaje escrito en actividades de aula que formen parte del proceso de enseñanza y aprendizaje. Proponen que el profesorado construya, en función de las características de la actividad, el contexto y su alumnado, una sencilla rejilla que contenga todos aquellos aspectos que responden a sus objetivos principales. Esta rejilla contiene por una parte los nombres de los alumnos y alumnas integrantes del grupo y, así mismo, una serie de indicadores que responden a los objetivos de cada actividad que se pretende evaluar. De esta forma, argumentan estas autoras, es posible ir evaluando de forma sistemática, continuada y personalizada, el trabajo de cada alumno. A modo de ejemplo, proponen un modelo como el siguiente:

Cuadro 4. Modelo de registro y evaluación de actividades de lenguaje escrito en el contexto de aula. (Tomado de Maruny, Ministral y Miralles, 2001).

E. Primaria

Texto expositivo

Lectura de definiciones sobre temas estudiados e identificación de los términos definidos.						
Lectura individual						
Alumnos/as	Niveles de lectura		Conocimientos previos acerca del tema	Comprensión del texto leído	Identificación de los términos de las definiciones	Justificación de las respuestas o hipótesis
	Interpreta a partir de índices textuales	Lectura ajustada al significado				

Desde este mismo modelo, otros autores como Jolibert (1984) o Repetto, Téllez y Beltrán (2002), proponen utilizar fichas o registros de *autoevaluación* que favorezcan, por una parte, una mayor implicación y participación activa del alumnado en el desarrollo de un elemento fundamental del proceso de enseñanza y aprendizaje como es la evaluación, y, por otra, una mayor toma de conciencia acerca del control y regulación del propio proceso de lectura y comprensión, así como acerca de los resultados obtenidos en el trabajo desarrollado. Un ejemplo de estas hojas de seguimiento y autoevaluación es el que Repetto, Sutil, Manzano, Téllez y Beltrán (2002) recogen en su programa de orientación metacognitiva *Comprender y Aprender en el aula*. En la ficha se recogen todas las unidades que el programa va desarrollando y un calendario en el que el alumno va anotando la puesta en práctica de las estrategias que son objeto de enseñanza.

Cuadro 5. Hoja de valoración y seguimiento por parte del propio alumno en el programa *Comprender y Aprender en el aula*. (Tomado de Repetto, Téllez y Beltrán, 2002).

		Curso _____ / _____																			
		Mes				Septiembre				Octubre				Noviembre				Diciembre..			
		Semanas																			
Estrategias	1. Identificar los objetivos de la lectura																				
	2. Activar los conocimientos previos																				
	3. Evaluar la tarea																				
	4. Planificar la lectura																				
	5. Desarrollar el vocabulario																				
	6. Identificar el tipo de texto																				
	7. Elaborar mapas conceptuales																				
	8. Detectar los significados del texto																				
	9. Autorregular la lectura																				
	10. Leer creativamente																				
	11. Leer y comprender rápido																				
	12. Hacer inferencias																				
	13. Hacer analogías																				
	14. Detectar las ambigüedades del texto																				
	15. Juzgar la lectura																				
	16. Detectar los errores de comprensión																				
	17. Solucionar esos errores																				
	18. Resumir el texto																				
	19. Hacer abstracciones																				
Progreso de cada mes																					

Sobre este tipo de evaluación que hemos descrito y en el que se destaca principalmente la contextualización de los instrumentos de evaluación y la consideración de ésta como parte integrante y fundamental del proceso de enseñanza y aprendizaje, volveremos más adelante, cuando exponamos y analicemos nuestra propuesta de evaluación dinámica de los procesos implicados en la lectura. Por el momento vamos a destacar algunas de las técnicas más específicas de evaluación cualitativa que se han ido desarrollando, algunas de las cuales han sido utilizadas en nuestro dispositivo. Esta valoración cualitativa, como hemos visto, ha ido muy de la mano de un concepto de evaluación como parte integrante del proceso educativo y que trata de obtener información de forma *continua* y *contextualizada* sobre éste, con el fin de regularlo y optimizarlo. De este modo, desde el comienzo, este tipo de evaluación se ha identificado con una evaluación del *proceso* de lectura (Repetto, Téllez y Beltrán, 2002), aunque también, como recogen los autores citados anteriormente, se han empleado algunas de estas técnicas de corte cualitativo para evaluar el producto de la comprensión, es decir: *los cambios de conocimiento e integración de nueva información que realiza el sujeto* (op. cit. pág. 251). Igualmente podemos decir que las estrategias metodológicas de evaluación cualitativa que han tratado de evaluar el proceso de lectura y comprensión, han estado estrechamente unidas a perspectivas que, de uno u otro modo, acentuaban la relevancia de los procesos de control y regulación metacognitiva, junto a un mayor grado de conciencia acerca de los propios conocimientos y de todo el proceso.

Técnicas y estrategias específicas de evaluación del proceso de comprensión

En cuanto a estas técnicas y estrategias de evaluación del proceso de lectura o de los distintos procesos implicados en ella, podemos destacar entre las que más se han utilizado: la entrevista (metacognitiva), el registro del pensamiento en voz alta, el análisis de la acción tutorial del alumno, la observación, la detección de incoherencias lingüísticas y las medidas directas del proceso de ejecución. A continuación vamos a analizar brevemente cada una de estas técnicas y estrategias cualitativas con el fin de seguir completando nuestra representación situacional del marco conceptual y metodológico relativo a la evaluación de la comprensión.

La entrevista

Denominada en ocasiones *entrevista metacognitiva*, se trata de una estrategia para acceder a lo que el alumno sabe sobre sus propios conocimientos y procedimientos para leer y comprender un texto, así como para acceder al modo en que tiene lugar el proceso mismo de comprensión, mediante la solicitud de su respuesta verbalizada. Se han distinguido dos tipos de entrevista en función del momento en el cual el evaluador propone la actividad, aunque también pueden diferir en el grado de estructuración de la misma. En un primer tipo, realizada antes o durante el proceso y denominada habitualmente *entrevista estructurada* o *semiestructurada* en función del grado de flexibilidad de las cuestiones a realizar, se pretende acceder al conocimiento del alumno sobre sus estrategias planteándole preguntas previas a la lectura y durante el proceso mismo de la lectura. El segundo tipo de entrevista al que se hace referencia en la literatura es la *entrevista retrospectiva*. En este caso se pide al alumno, una vez leído el texto, que exprese cómo ha realizado la tarea de lectura determinada y qué estaba pensando mientras estaba leyendo. Se trata de indagar en los procesos de supervisión que ha puesto en marcha el alumno, así como en su metaconocimiento estratégico.

Las principales limitaciones descritas en la utilización de la entrevista hacen referencia a los problemas que plantea la accesibilidad de los procesos cognitivos a la conciencia, sobre todo cuando hablamos de alumnos y alumnas de corta edad (Flavell, 1992; Garner y Alexander, 1989). En efecto, algunas investigaciones han informado de que el desarrollo del conocimiento metacognitivo tiene lugar de forma tardía con respecto a la posibilidad de proceder de forma estratégica, lo que podría explicar en parte el hecho de que se manifieste la puesta en práctica de estrategias de control y regulación cognitiva y en cambio sea difícil acceder conscientemente a la información sobre la propia ejecución (Brown, 1987; Alexander et. Al., 1995; Navarro y Mora, 2002). Otras limitaciones que se han descrito, relacionadas en parte con las anteriores, aluden a que el alumno puede conocer lo que debe hacer, es decir, puede poseer el conocimiento adecuado sobre una estrategia, pero en cambio no ponerla en práctica. Puede suceder también lo contrario, como apuntábamos; es decir, que el alumno sea capaz de poner en práctica estrategias de lectura, pero no sea capaz de verbalizar cómo lo ha hecho. La causa de todo esto podría estar en que la capacidad de expresión verbal del alumno está interfiriendo en sus respuestas verbales, lo que explicaría que éstas fueran erróneas aunque de la ejecución de una determinada actividad pudiera inferirse la puesta en práctica de estrategias (Brown, 1980; Cavanaugh y Perlmutter, 1982; Ericson y Simon, 1980; Nisbett y Wilson, 1977). Otra posibilidad puede tener que ver con la dificultad que supone llevar a la conciencia procesos que podrían estar muy automatizados (Brown, 1987; Garner, 1987); en este caso, los mayores problemas se encontrarían entre

adultos o alumnos ya mayores que han logrado interiorizar y automatizar determinados procedimientos hasta el punto de encontrar serias dificultades para poder acceder de forma consciente a ellos. Brown también destaca que un posible problema es que algunas estrategias de control y regulación cognitiva son aprendidas sin que haya mediado una reflexión consciente en el aprendizaje, lo que conllevaría cierta dificultad para informar verbalmente a otros sobre estos procesos. Por último, otra posible limitación de esta estrategia consiste en que las respuestas del alumno puedan estar guiadas más por lo que cree que debe responder, en función de lo que piensa que el evaluador espera escuchar, que por una reflexión sobre sus propios pensamientos (Meichenbaum et al., 1985). Como tuvimos oportunidad de valorar en otro trabajo (Navarro y Mora, 2002), en el que nos adentrábamos en los estudios que han abordado la evaluación de la metacognición en la comprensión de textos, y más concretamente las actividades específicas que han sido utilizadas en este tipo de evaluación; autores como Meichenbaum (1985) o Garner (1987) han tratado de superar algunas de las limitaciones metodológicas que conlleva la puesta en práctica de instrumentos como la entrevista, aportando algunos elementos a modo de requisitos que deberían ser considerados previamente para maximizar la validez de la información recogida en este tipo de actividades.

El registro del pensamiento en voz alta

Esta estrategia consiste en pedir al alumno que nos cuente, mientras está leyendo, todo lo que piensa con relación a la lectura; se pretende acceder de esta forma a los procesos de control y regulación cognitiva de la lectura que el alumno desarrolla, ya que puede informarnos —directa o indirectamente— sobre la puesta en práctica de estrategias de superación de las dificultades encontradas, de control de la impulsividad, de estrategias relacionadas con el control de la atención, el uso de la mediación verbal, el control de la actividad para determinar si entiende o no lo que está leyendo o sobre la consideración del grado de dificultad o interés que le suscita la lectura. También podemos obtener información sobre la realización de inferencias, la formulación de hipótesis y predicciones o la activación de conocimientos previos y su conexión con los contenidos del texto. Las limitaciones de esta estrategia de evaluación están ligadas igualmente a la capacidad de expresión verbal del sujeto y a las dificultades de llevar a la conciencia procesos cognitivos y metacognitivos; también se ha hecho referencia a las posibles interferencias que la intervención del evaluador durante el proceso puede causar en las respuestas del alumno, así como a la posible interferencia del proceso natural de pensamiento debido a que éste debe verbalizarse durante el proceso (Alonso Tapia et al., 1992; Moreno y García, 2000; Repetto, Téllez y Beltrán, 2002).

La observación

La observación como estrategia de recogida de información consiste, en palabras de Fernández-Ballesteros (1980), en «*una conducta deliberada del observador, cuyos objetivos van en la línea de recoger datos en base a los cuales poder formular o verificar hipótesis*». Se trataría, por tanto, de una observación intencionada, que puede ser en mayor o menor medida sistemática y estructurada, teniendo por objeto en este caso las conductas (cognitivas) y la actitud del alumno con relación a los textos y la lectura.

Siguiendo a esta misma autora (Fernández-Ballesteros, 1992), la observación puede realizarse en contextos naturales o artificiales. La escuela y el trabajo desarrollado en las aulas podría ser considerado un contexto natural; siendo un buen ejemplo de observación en este contexto, realizada para extraer información conducente a evaluar el proceso de enseñanza y aprendizaje relacionado con el lenguaje escrito, el que recogíamos anteriormente cuando exponíamos la evaluación de la lectura desde una concepción constructivista, haciendo referencia a los trabajos y las propuestas de Bofarull (2001) o Maruny, Ministral y Miralles (2001). De otro modo, pueden introducirse también deliberadamente ciertas modificaciones en el contexto natural del trabajo desarrollado en un aula, con el objetivo de observar determinadas conductas en el alumnado (Moreno, 1996); de esta forma, la introducción de actividades específicas de evaluación, como puede ser una prueba objetiva que no forma parte habitual del proceso o que es realizada por personas ajenas al contexto escolar, puede transformar al menos en parte el contexto natural en ciertamente artificial, teniendo lógicamente consecuencias relacionadas con la interpretación o la validez ecológica de los resultados obtenidos (Bronfenbrenner (1987).

Los tipos de registro utilizados para recoger datos provenientes de la observación suelen tener formatos variados, que a menudo están condicionados por la información que se desea registrar y/o por la metodología considerada idónea para hacerlo. Básicamente podemos encontrar dos tipos: registros de tipo narrativo y sistemas de registros categoriales (Lukas y Santiago, 2004). Aunque su distinción en la práctica investigativa, y sobre todo educativa, no deja de ser un tanto artificial, ya que muchos de los instrumentos de observación utilizados contienen elementos que caracterizan ambos tipos de registro, recogemos a continuación sus características esenciales, lo que nos servirá también posteriormente, en la segunda parte de este trabajo, cuando nos centremos en nuestra propuesta de evaluación.

En los *registros narrativos* el observador anota comportamientos y sucesos que tienen lugar en un escenario determinado. Estas anotaciones no suelen estar limitadas a conductas observables, sino que en este tipo de registros las anotaciones pueden contener inferencias del evaluador con base en la conducta observable de los sujetos, cobrando éste un protagonismo esencial en el proceso. La formación del evaluador y, en ocasiones, la determinación previa y precisa de los datos a observar suele ser también de enorme importancia en la práctica en este proceso de observación. Los registros narrativos más utilizados incluyen (a) las *notas de campo*, donde un observador que participa activamente de una actividad registra unos determinados comportamientos y sucesos, así como su interpretación. Por ejemplo, un profesor que evalúa la puesta en práctica de estrategias de comprensión por parte de su grupo de alumnos, anotando lo que estos hacen o dicen con respecto a las actividades propuestas y extrayendo de su comportamiento información relativa a sus habilidades, estaría utilizando este tipo de registro. Otros registros narrativos comúnmente utilizados en la escuela, o propuestos en investigación educativa, son (b) los *diarios*; en ellos el profesor puede anotar incidencias, comportamientos y pensamientos relativos a su práctica docente a lo largo de un determinado tiempo. Un tercer tipo de registros narrativos lo constituyen los *registros anecdóticos*, en los que se recogen y describen de forma puntual una serie de hechos o comportamientos de especial relevancia. Por último, (d) *la descripción de muestras de conducta* hace referencia al registro de comportamientos o sucesos por parte de un observador que no participa activamente de la actividad. Este tipo de registro también es utilizado en propuestas de evaluación e intervención educativas en

las que un profesor/evaluador observa, registra e interpreta unos determinados comportamientos que acontecen en un aula (Parrilla, 1996; Parrilla y Daniels, 1998).

Por otra parte, los sistemas de *registros categoriales* suponen la determinación previa y precisa de los comportamientos o hechos que deben ser observados. Consisten, por tanto, en la anotación de observaciones con base en un sistema de categorías previamente determinado. Pueden distinguirse dos grandes grupos: las *escalas valorativas* y la *lista de control*. En el primero el observador debe anotar la frecuencia e intensidad con que aparecen unas determinadas conductas o sucesos que forman parte de categorías previamente establecidas. En el segundo grupo, el observador debe fijar su atención en la presencia o ausencia de una determinada conducta o suceso previamente fijado y categorizado. Como apuntábamos anteriormente, estas agrupaciones pueden verse limitadas en la práctica por la utilización de instrumentos de observación que contienen categorías prefijadas de conductas a observar (estrategias de lectura) y que, al mismo tiempo, proponen el registro narrativo e interpretativo de un observador participante sobre su puesta en práctica (Parrilla, 1996; Bofarull, 2001; Maruny, Ministral y Miralles, 2001; Navarro y Mora, 2003b).

En los últimos años se ha retomado el interés de forma significativa por los métodos observacionales en psicología y educación. A menudo esta técnica es complemento de otras como la entrevista o la pasación de pruebas objetivas. Una extensión de las posibilidades de la observación como estrategia de evaluación en el contexto escolar es la utilización del video. Como destacan, entre otros, Mora (1998) o Moreno (1996), las ventajas de su uso estriban principalmente en la posibilidad de realizar micro-análisis de las conductas acaecidas (por ejemplo, de las interacciones que mantienen una madre con su hijo o una profesora con un grupo de alumnos) y en la posibilidad de analizar la información repetidas veces y por un mayor número de personas, incrementando de este modo la información recogida y permitiendo, en el campo de la investigación, análisis que conduzcan a una mayor fiabilidad de las observaciones e interpretaciones realizadas.

En cuanto a sus limitaciones, éstas están relacionadas con tres factores: el sujeto o grupo de sujetos observado, el observador y el sistema de observación empleado (Fernández Ballesteros, 1992). En efecto, en situaciones y contextos «artificiales», la conducta de los alumnos puede experimentar cambios sustanciales que no reflejarían su comportamiento habitual en situaciones más naturales; es lo que se denomina *reactividad*, la cual representa un importante hándicap a la hora de interpretar y generalizar los resultados. Por lo que respecta al observador, los posibles sesgos en la información recogida proceden básicamente (a) del grado de participación que éste tenga, estableciéndose hasta tres grados: observador no participante (p.e., el profesor que recoge información sobre las conductas de relación entre iguales que acontecen durante el recreo), observador ajeno/participante (p.e., el orientador que realiza una valoración psicopedagógica de un alumno con necesidades educativas especiales) y observador allegado/participante (p.e. el profesor habitual de un grupo de alumnos que evalúa su competencia en lectura en el transcurso de las actividades de enseñanza y aprendizaje); (b) de las expectativas que se generan en el observador acerca de la situación evaluada, lo que se pone de manifiesto, como recoge Rosenthal (1966, 1976: citado en Fernández Ballesteros, 1992) en distintos momentos del proceso, desde la elección del sistema de observación, hasta el registro de las conductas en función de las hipótesis previas, pasando por la elección del proceso estadístico y de elaboración de los

datos observados; y (c) del entrenamiento y preparación del observador. Por último, están las limitaciones achacables al propio sistema de observación utilizado. Éstas serán mayores cuando no estén claramente definidas, o no sean bien conocidas por el observador la situación o las conductas a observar. También puede ocurrir que el número de categorías o conductas que se pretenden observar sea muy elevado, influyendo negativamente en la fiabilidad y la validez de las observaciones (Kent y Foster, 1977).

La tutorización entre iguales

Consiste en pedir a un alumno que ayude a un compañero —normalmente menos competente o de menor edad— en tareas de lectura y comprensión. Se le puede pedir que apoye a su compañero en la realización de una tarea determinada, lo que nos informaría sobre su conocimiento de las estrategias que considera adecuadas para ayudar de forma eficaz y resolver la tarea, así como del procedimiento que sugiere para su realización. Es importante que el alumno-tutor no realice la actividad del compañero, sino que trate de ayudarle ofreciéndole sugerencias sobre cómo puede solucionar la actividad planteada (Mora, 1998; Repetto, Téllez y Beltrán, 2002). De otro modo, inicialmente, también puede pedírsele que explique a un compañero cómo se realiza una determinada actividad de lectura. El análisis de sus explicaciones, lógicamente teniendo en cuenta las limitaciones relacionadas con la capacidad de expresión verbal que pueda tener el alumno y que hemos descrito con anterioridad, podría informarnos sobre su metaconocimiento estratégico.

Detección de incoherencias lingüísticas

Esta estrategia consiste en la introducción intencional de erratas o incoherencias lingüísticas a nivel léxico, sintáctico o semántico en un texto que es posteriormente presentado al alumno sin advertirle de la existencia de estos errores. La detección supondría la puesta en práctica de mecanismos de supervisión de la propia comprensión (Alonso Tapia et. al., 1992; Moreno y García, 2000; Repetto, Téllez y Beltrán, 2002; Navarro y Mora, 2003a), y las acciones encaminadas a dotar de sentido el texto podrían desvelar, además del uso de estrategias de supervisión y autoevaluación, conocimientos de la estructura sintáctica y de la estructura textual, así como la activación e integración de conocimientos previos. Algunas limitaciones recogidas por Alonso Tapia y sus colaboradores (1992) con respecto a esta actividad hacen referencia a que: (a) los niños tienden a pensar que los textos que se les proporciona desde la escuela están bien escritos y por tanto no están preparados para detectar inconsistencias o errores; (b) la realización de inferencias por parte del alumno podría llevar a dar coherencia a la organización y al contenido del mensaje, aunque no lo tenga; (c) el hecho de que la unidad de medida empleada comúnmente sea la verbalización del alumno puede introducir sesgos relacionados con la interferencia entre la conducta y la expresión verbal del alumno; (d) la naturaleza del contenido de los textos puede influir notablemente en el grado de detección de las incoherencias (Garner y Anderson, 1982: citado en Alonso Tapia et. al., 1992); y (e) al hecho de que las aptitudes de tipo verbal puedan influir en la detección, interfiriendo de esta forma con la propia capacidad de supervisión.

Medidas directas del proceso de comprensión

Se trataría de la utilización de unidades de análisis conductuales y, por tanto, observables directamente, aunque en la práctica algunas de ellas requieren de interpretación por parte del observador para que la información obtenida sea de utilidad. Algunas medidas de este tipo proceden de la psicofisiología o guardan una estrecha relación con ésta; por ejemplo, el análisis de los movimientos oculares durante la lectura (fijaciones, velocidad y amplitud de los movimientos, regresiones, etc.). Repetto, Téllez y Beltrán (2002) recogen otras medidas directas del proceso de lectura como por ejemplo las mediciones del tiempo de lectura, ciertos gestos que delatan la detección de un problema o el análisis de errores cometidos en la lectura en voz alta. En cuanto a esta última medida mencionada, Alonso Tapia y sus colaboradores hacen referencia a que el análisis de la comisión de errores en la lectura en voz alta y su corrección espontánea se relacionaría con el uso que hacemos de señales sintácticas, semánticas y grafofónicas durante la lectura, y que los errores cometidos, así como las correcciones efectuadas mostrarían qué tipo de procesamiento se está llevando a cabo. Por último, Repetto y su grupo, citando a August, Flavell y Clift (1984) y a Baker y Brown (1984), hacen referencia al análisis de la interacción que mantiene el sujeto con el ordenador en la realización de actividades de lectura. En este caso, las medidas pueden consistir en una serie de conductas que el ordenador registra mientras el alumno realiza una actividad, por ejemplo, volver atrás para leer de nuevo una palabra o un párrafo, la velocidad de lectura o el seguimiento correcto de una instrucción dada. Estas medidas requieren necesariamente de interpretación por parte del evaluador.

Cobra especial importancia, por una parte, la fiabilidad del registro de los datos provenientes de este tipo de medidas, así como la concreción de las conductas que deben ser observadas y registradas. En este sentido, una de las limitaciones más importantes de este tipo de análisis consiste en que en la evaluación del proceso de enseñanza y aprendizaje de la lectura nos puede interesar de forma significativa recoger información sobre conductas que no son observables directamente, siendo en estos casos necesario inferir o interpretar determinados comportamientos observables poniéndolos en relación con la puesta en práctica de procesos cognitivos o emocionales.

Técnicas y estrategias específicas de evaluación del producto de la comprensión

En cuanto a las técnicas o estrategias de evaluación del producto de la comprensión, destacamos entre las comúnmente utilizadas las de recuerdo libre y provocado, resumen, evaluación mediante preguntas sobre el texto y las pruebas de cierre o completamiento (Johnston, 1989; Repetto, Téllez y Beltrán, 2002; Alonso Tapia et. al., 1992).

Prueba de recuerdo libre

Esta técnica consiste en pedir al alumno, tras la lectura de un texto, que nos cuente con sus propias palabras el contenido de lo que ha leído. Es importante tener en cuenta que al tratarse de una tarea de recuerdo, los resultados podrían estar contaminados por la capacidad de memoria a corto y largo plazo, por lo que una posible forma de evitar o atenuar esto sería evaluar con el texto presente. Por otra parte, es relevante igualmente valorar el grado de integración textual y representación global del significado del texto construido por el alumno, así como la puesta en práctica de estrategias como la

realización de inferencias, la integración de conocimientos previos de forma adecuada para comprender y revisar el contenido desde lo que sabe, etc. Por su parte, Johnston (1989) recoge otros posibles problemas que pueden interferir en la respuesta del alumno sobre su comprensión; entre estos cita: (a) las dificultades de expresión verbal o escrita que tenga el sujeto; (b) la necesidad de que el lector sea consciente del propósito de la tarea, es decir, debe saber con claridad «qué nivel de detalles debe retener y reproducir»; (c) la dificultad que supone situarse en cierta forma en la perspectiva del evaluador, debiendo elaborar una hipótesis acerca de lo que éste espera; (d) el bloqueo que puede sufrir el alumno a la hora de recuperar la información; y (e) la posibilidad de que se hayan malinterpretado las demandas de la tarea, otorgando el sujeto de esta forma una respuesta inadecuada, que no sería reflejo de lo que se quería evaluar.

Recuerdo provocado

Se trata de una técnica que complementa a la de recuerdo libre, realizando preguntas al alumno que faciliten su acceso a la información sobre el texto, información que, tal vez no fue capaz de evocar y/o organizar de forma espontánea. Johnston (1989), citando a Anderson y Pichert (1977), señala algunas limitaciones relacionadas con el recuerdo provocado, parte de las cuales quedaban ya recogidas anteriormente, se relacionan con ellas o pueden ser aplicadas igualmente a la actividad de recuerdo libre. Entre estas limitaciones se encuentra el hecho de que en función de la perspectiva teórica o conceptual desde la que se realicen las preguntas, se podría obtener una información diferencial del alumno; otra limitación consiste en la dificultad de generar preguntas en función de cada texto que sean relevantes para obtener información valiosa sobre el proceso de comprensión o los distintos procesos implicados; un tercer problema lo constituye la dificultad de procesar, analizar y evaluar el contenido de las respuestas del alumnado en protocolos orales o escritos, tanto de tareas de recuerdo libre como de recuerdo provocado; por último, persiste el problema de separar las posibles interferencias de la capacidad de memoria en el proceso de comprensión.

Resumen

Con esta actividad se pretende valorar la capacidad del alumno para elaborar una síntesis significativa del contenido de un texto. Se solicita que el alumno realice un resumen de la información leída de forma que recoja los aspectos relevantes de la misma y muestre haber construido una representación global del significado del texto. Se trata de una tarea, en esencia, similar a las anteriores. Puede ser igualmente realizada oralmente o por escrito, teniendo en cuenta que los contenidos destacados por un alumno en su composición escrita pueden reflejar no solo su capacidad de síntesis y la representación que del contenido del texto ha elaborado, sino también sus habilidades de expresión escrita. En este sentido, para evitar este problema, Alonso Tapia, Carriedo y González (1992), en la construcción de la batería IDEPA para la *evaluación de la capacidad de comprender y resumir lo importante*, proponen la sustitución de la actividad de composición escrita de un resumen por la de identificar entre varios el párrafo que mejor resume el texto leído. Por otra parte, al igual que anteriormente para la actividades de recuerdo libre o provocado, se plantea la alternativa de la disponibilidad del texto durante la tarea de resumen para evitar la interferencia de la capacidad de memoria en los resultados de la actividad (Johnston, 1989; Alonso Tapia, Carriedo y González, 1992; Català, Català, Molina y Monclús, 2001).

Evaluación mediante preguntas sobre el texto

Se trata de la actividad de evaluación de la comprensión más comúnmente utilizada en la educación escolar. Aunque presenta variedad de formatos y también de propósitos, básicamente, la mayoría de las actividades o pruebas de evaluación elaboradas buscan recoger información sobre tres aspectos: los conocimientos que ha adquirido el alumno con base en un texto leído, la representación textual que ha construido el alumno, y la integración de nueva información que ha realizado. Por otra parte, las preguntas —orales o escritas— pueden estar formuladas de forma que se requieran para su contestación distintos niveles de procesamiento. A continuación recogemos distintos tipos de preguntas que componen actividades o pruebas de evaluación del producto de la comprensión. Como comprobará el lector, la secuencia descendente que seguimos no solo se corresponde con el nivel de procesamiento exigido al alumno para la respuesta, sino que también se relaciona en parte con la secuencia histórica que ha seguido la confección de este tipo de actividades de evaluación con respecto al contenido de las preguntas formuladas. De este modo, determinadas preguntas sobre un texto pueden requerir una búsqueda o rastreo sobre la base textual, dentro de la cual puede hallarse más o menos literalmente la respuesta; el grado o nivel de procesamiento sería en este caso superficial y la comprensión requerida podría ser denominada literal (Català, Català, Molina y Monclús, 2001; Repetto, Sutil, Manzano, Téllez y Beltrán, 2002); de otro modo, otras preguntas pueden plantear grados de procesamiento menos superficiales e ir dirigidas a que el alumnado realice una reorganización de las ideas del texto, pudiendo efectuar operaciones de clasificación, esquematización, resumen o síntesis del contenido leído. Un tercer grupo de preguntas que puede formar parte de actividades o pruebas de evaluación de la comprensión, puede solicitar del alumno la elaboración de una comprensión inferencial o interpretativa del texto; en este caso, el grado de procesamiento es más profundo y el alumno debe realizar actividades como deducir ideas principales contenidas en el texto, establecer relaciones de causa-efecto, conexiones no explícitas entre ideas, predecir resultados o acontecimientos, formular hipótesis ,etc. Un cuarto grupo de preguntas que pueden formularse van encaminadas a evaluar la comprensión crítica del alumno (Alonso Tapia y Gutiérrez, 1992); los requerimientos en este caso se dirigen a que elabore juicios sobre hechos, opiniones, validez de un determinado elemento, realidad o fantasía, conveniencia, etc. Otro tipo de cuestiones con profundo grado de procesamiento requerido sería el que se dirige a la comprensión de la intencionalidad del autor o autora, es decir, a la capacidad manifestada para sacar conclusiones sobre los razonamientos esgrimidos en un texto argumentativo, captar intenciones o sentimientos expresados de forma indirecta, etc. (Català, Català, Molina y Monclús, 2001). Por último, queremos hacer alusión a un posible grupo de preguntas dirigidas esencialmente a valorar el conocimiento metacognitivo del alumno con relación a las estrategias de comprensión, así como la puesta en práctica de estrategias de autorregulación en el proceso de comprensión. Estas preguntas requieren que el alumno efectúe una reflexión sobre sus propios conocimientos y estrategias para comprender, o bien que realice operaciones de supervisión y evaluación del propio proceso de comprensión, requiriendo por tanto un alto nivel de procesamiento (Alonso Tapia, Carriedo y Mateos, 1992; Alonso Tapia, Carriedo y González, 1992; Peronard, Velásquez, Crespo y Viramonte, 2002; Navarro y Mora, 2003a).

Podemos, así mismo, distinguir entre actividades que plantean *preguntas abiertas* o *cerradas*. Habitualmente, las primeras poseen un grado de sistematización y

estructuración menor que las segundas, aunque existe una tendencia que despierta especialmente nuestro interés, consistente en dotar del máximo nivel posible de estas características de sistematización y estructuración a cuestiones planteadas en un formato abierto y que a su vez, contiene elementos que difuminan los límites con la evaluación del proceso de lectura, planteando la intervención del evaluador durante el transcurso de la evaluación. Dentro de este primer grupo de actividades que proponen cuestiones abiertas, podríamos por tanto diferenciar (a) los *Cuestionarios acumulativos* (Repetto, Téllez y Beltrán, 2002), que recogen una amplia gama de preguntas acerca de distintos elementos relacionados con la lectura realizada: tema, idea principal, identificación de la intencionalidad del autor, relación entre contenidos, etc.; y (b) *los Cuestionarios sistemáticos de evaluación*, actividad a la que hacíamos referencia anteriormente. Con este nombre, algunos autores como Alonso Tapia (1995) o Carriedo y Alonso (1994), designan un tipo de evaluación que se aproxima teórica y metodológicamente a la evaluación dinámica o interactiva y que, así mismo, comparte elementos con la entrevista metacognitiva (retrospectiva) descrita en páginas anteriores. Se propone indagar en las respuestas dadas por el alumno, profundizando en ellas con nuevas preguntas que buscan la justificación de las respuestas, pudiendo acceder de esta forma a las estrategias que el alumno ha utilizado para comprender y a los conocimientos que posee acerca del texto o de sus propias estrategias. Esta forma de evaluación, como podemos comprobar, permite en cierta medida analizar el proceso de comprensión, la calidad del metaconocimiento estratégico y el uso de las estrategias de lectura; sin embargo, es igualmente importante destacar que no es suficiente con que una determinada técnica o estrategia de evaluación nos de la posibilidad de evaluar procesos y de hacerlo de forma interactiva para obtener información valiosa para la acción educativa, sino que posee la misma importancia el propósito con el que se aplica esta técnica de evaluación, de forma que si ésta es considerada en su aplicación como una fuente de información sobre aptitudes ya cristalizadas y no en continuo desarrollo, la información obtenida y analizada desde esa óptica por los evaluadores tendrá poco que ver con la que se efectuaría desde una perspectiva dinámica.

El segundo grupo de actividades al que hacíamos referencia es el que plantea preguntas cerradas. Este tipo de actividades consiste en la presentación de preguntas, habitualmente tras la lectura de un texto, a las que acompañan varias respuestas entre las que se encuentra la considerada como correcta o verdadera. Existen básicamente dos modalidades: las pruebas que recogen *preguntas con opción múltiple de respuesta* y las *pruebas de verdadero o falso*. En las primeras se presentan varias respuestas al alumno entre las que éste tiene que elegir la correcta; habitualmente solo una respuesta es la correcta y el número de opciones oscila entre tres y cinco, aportando mayor validez un número de opciones mayor, ya que minimiza la intervención del azar en la elección de respuesta (Català, Català, Molina y Monclús, 2001). Así mismo, para que estas pruebas permitan obtener información valiosa sobre las distintas variables intervinientes en el proceso de comprensión, o incluso para informar sobre determinados procesos implicados en la comprensión a través de las elecciones realizadas, es de suma importancia que la construcción de la prueba permita una cierta aproximación cualitativa con base en las opciones marcadas por el alumno, tanto aquellas que representan aciertos, como las que indican errores (Repetto, Téllez y Beltrán, 2002; Alonso Tapia, Carriedo y Mateos, 1992; Català, Català, Molina y Monclús, 2001; Peronard, Velásquez, Crespo y Viramonte, 2002; Navarro y Mora, 2002).

Por otra parte, se trata de pruebas que, bien construidas, con base en una teoría de la enseñanza y el aprendizaje de la lectura, en función de los procesos que intervienen en la misma y atendiendo a unas adecuadas características técnicas en cuanto a las tipologías y el contenido de los textos elegidos, así como en cuanto a la presentación y redacción de los ítems de respuesta que reflejan aciertos o errores, pueden llegar a otorgar una información útil para la evaluación que, como señalan Alonso Tapia y sus colaboradores (1992), deberá ser, en cualquier caso, contrastada por otros métodos. Permiten, así mismo, una codificación objetiva, minimizando los posibles aciertos al azar que tienen las pruebas de verdadero o falso que veremos a continuación. Por otra parte, entre las limitaciones reconocidas a este tipo de pruebas estarían (a) la posibilidad de que se entienda el texto presentado y no así la pregunta o las opciones formuladas (Català, Català, Molina y Monclús, 2001); y, lógicamente, (b) la imposibilidad de acceder al proceso interactivo y dinámico que supone leer y comprender.

Las preguntas de verdadero o falso suponen una alternativa que, planteada con base en solo dos opciones de respuesta, plantea diversas limitaciones. Unida a la expuesta anteriormente relativa al problema de las respuestas dadas al azar (Repetto, Téllez y Beltrán, 2002), otros autores plantean el hecho de que no podemos conocer ni interpretar adecuadamente los motivos últimos de la contestación debido, por una parte, al elevado porcentaje de posibilidades de acierto y también a la confusión que puede tener lugar en lectores con baja competencia entre la estructura superficial del texto y de las opciones y el contenido de las mismas, pudiendo elegir el alumno una opción equivocada con base en la estructura superficial de la misma, incluso aunque haya comprendido el texto; esta confusión puede estar originada en muchos casos sobre la base de aspectos relacionados con una autopercepción negativa (Johnston, 1989; Carriedo y Alonso, 1994; Alonso Tapia, 1995; Marín, 1994).

Un intento de evitar las limitaciones de las pruebas de verdadero o falso y, al mismo tiempo, conservar la esencia de este tipo de preguntas, ha consistido en introducir en pruebas de opción múltiple cuestiones que plantean al alumno la necesidad de elegir qué opción de las que se proponen es verdadera o cuál es falsa (Català, Català, Molina, Monclús, 2001).

Pruebas de cierre o completamiento

Consiste en la presentación al alumno de un texto en el que faltan palabras, solicitándole que complete el mensaje. Se trata de las conocidas pruebas tipo CLOZE (Suárez y Meara, 1992). Permiten evaluar la capacidad de establecer relaciones de tipo sintáctico y semántico principalmente, ofreciendo información —en función de la construcción del texto y de la relevancia y categorías funcionales de las palabras ausentes— sobre la comprensión, debiendo hacer uso el alumno del contexto para completar el texto. El uso de estas pruebas ha estado habitualmente vinculado a la valoración del producto de la comprensión; sin embargo, como recogen Johnston (1989) o Alonso Tapia, Carriedo y Mateos (1992), es una técnica que puede ofrecer alguna información sobre el proceso de lectura, dado que el completamiento del texto, como hemos recogido anteriormente, requeriría el uso del contexto, un aspecto relevante del proceso de comprensión. Alonso Tapia y sus colaboradores destacan una de las principales limitaciones de esta prueba; ésta consiste en que al otorgar de antemano inconsistencias o lagunas en el texto —las palabras que faltan, sustituidas por espacios en blanco— se reducen o eliminan las posibilidades de evaluar de forma natural la supervisión y regulación de la comprensión

lectora, aspectos que inciden notablemente en la comprensión de muchos lectores, no pudiéndose de esta forma obtener un indicador adecuado de la capacidad de comprensión del sujeto en un texto sin lagunas. Johnston (1989), por su parte, destaca igualmente que las demandas de la tarea no se corresponden con las de la lectura natural de un texto, por cuanto es necesario mantener en la memoria inmediata la información procedente de la proposición en la cual se encuentra el espacio en blanco hasta poder completar la oración una vez localizada la información ausente.

Criterios de evaluación utilizados en la evaluación de la lectura y la comprensión

La investigación de un diverso grupo de psicólogos educativos y de otros profesionales relacionados con los procesos de enseñanza y aprendizaje, que han dedicado su esfuerzo y su trabajo durante las últimas décadas al análisis y estudio de la lectura, ha ido generando y conformando un cierto cuerpo de conocimientos acerca del sentido y significado del hecho de leer, de los procesos y elementos que intervienen en la comprensión de un texto y en el aprendizaje a partir de estos, así como de las estrategias que son necesarias para acceder a una representación global del significado que permita comprender, aprender, analizar y ampliar el contenido textual.

Por otra parte, como veíamos con anterioridad, la relevancia adquirida por la lectura como herramienta fundamental en el acceso continuo al conocimiento y en el análisis de la realidad social; así como por la enseñanza de la lectura desde una perspectiva que sea acorde con el planteamiento anterior, ha contribuido a la progresiva integración de las aportaciones realizadas desde la investigación sobre la lectura y la comprensión en los procesos de enseñanza y aprendizaje de la lectura en la escuela; esto ha llevado lógicamente a la necesidad de establecer unos criterios de evaluación que respondieran a los contenidos que integran el trabajo para enseñar la lectura y la comprensión. Se han propuesto distintos conjuntos de criterios de evaluación en función de la óptica teórica de base adoptada. Recogemos a continuación (cuadro 6) algunos de estos criterios junto con referencias de autores que, o bien los incluyen en sus propuestas, o bien destacan su relevancia en el proceso de evaluación.

Cuadro 6. Criterios de evaluación de la lectura y la comprensión de textos que han sido destacados por diversos autores e incluidos en sus propuestas de evaluación

<i>Aspectos perceptivo-visuales</i>	Filho, 1973; Mora, 1993; Defior y Ortuzar, 1993; Cuetos, Rodríguez y Ruano, 1996; De la Cruz, 2000; Romero y Lavigne, 2006
<i>Habilidades metalingüísticas/conciencia fonológica</i> <ul style="list-style-type: none"> • Conciencia fonémica y silábica • Conciencia léxica • Conciencia sintáctica • Conciencia semántica 	Manis y Morrison, 1985; Lundberg, Frost y Petersen, 1988; Lundberg y Høien, 1991; Carrillo, 1993; Ortiz, 1994; Wagner, Torgesen y Rashotte, 1994; Jiménez y Ortiz, 1998; Jiménez, Gregg y Díaz, 2004

<i>Descodificación</i>	Toro y Cervera, 1984; González Portal, 1985; Maldonado, Sebastián y Soto, 1992; Defior y Ortuzar, 1993; Alonso Tapia y Carriedo, 1996; Cuetos, Rodríguez y Ruano, 1996; Ramos y Cuetos, 2000; Romero y Lavigne, 2006
<i>Lectura en voz alta y velocidad lectora</i>	Solè, 1992; Bofarull, 2001; González Portal, 1985; Català, Català, Molina y Monclús, 2001
<i>Evaluación de la memoria y de estrategias de memoria</i>	Filho, 1973; Johnston, 1989; Defior y Ortuzar, 1993; Mora, 1993; De la Cruz, 2000; Kirk, McCarthy y Kirk, 1928/2004; Swanson, 1992, 2000; Romero y Lavigne, 2006
<i>Actitud emocional e interés con relación a la lectura</i>	Solè, 1992; Bofarull, 2001; Mora Merchán y Mora, 2000; Maruny, Ministral y Mirallés, 2001; Navarro y Mora, 2002; Romero y Lavigne, 2006
<i>Evaluación del contexto (proceso de enseñanza y aprendizaje)</i> <ul style="list-style-type: none"> • Diseño y desarrollo de las actividades de lectura • Ideas de los profesores sobre el proceso de E-A de la lectura • Uso del espacio, tiempo y recursos con relación a la enseñanza de la lectura • Dinámica y clima del aula (estrategias docentes de interacción con el alumnado y pautas de interacción entre compañeros). • Indicadores de progreso en lectura • El sistema de evaluación adoptado 	Mirás y Solè, 1990; Moreno y García, 2000; Bofarull, 2001; Maruny, Ministral y Mirallés, 2001; Mora, 1997; Romero y Lavigne, 2006
<i>Evaluación de los textos/tareas</i> <ul style="list-style-type: none"> • Progresión temática • Estructuración del texto • Significado global • Ayudas o claves textuales • Coherencia textual 	Moreno y García, 2000; Sánchez, 1993; Aznar, Cross y Quintana, 1993; Romero y Lavigne, 2006
<i>Concepción sobre la lectura</i>	Repetto, 1997; Repetto, Téllez y Beltrán, 2002; Alonso Tapia y Carriedo, 1996; Maruny, Ministral y Mirallés, 2001; Martín, 1999; Moreno y García, 2000; Blanton, Wood y Moorman, 1990; Alonso Tapia, Carriedo y Mateos, 1992; Navarro y Mora, 2003a

<p><i>Objetivo/Propósito de la lectura</i></p>	<p>Repetto, 1997; Repetto, Téllez y Beltrán, 2002; Maruny, Ministrál y Mirallés, 2001; Bofarull, 2001; Solè, 1992; Navarro y Mora, 2003a</p>
<p><i>Conocimiento/Metaconocimiento sobre estrategias de comprensión</i></p> <ul style="list-style-type: none"> • Sobre qué estrategias utilizar • Cómo desarrollarlas • Cuándo ponerlas en práctica 	<p>Alonso Tapia, Carriedo y Mateos, 1992; Alonso Tapia, Carriedo y González, 1992; Alonso Tapia y Carriedo, 1996; Johnston, 1989; Moreno y García, 2000; Peronard, Velásquez, Crespo y Viramonte, 2002; Navarro y Mora, 2003a</p>
<p><i>Evaluación de los conocimientos previos del lector</i></p> <ul style="list-style-type: none"> • Vocabulario • Conocimientos que posee sobre el tema • Organización de sus conocimientos • Esquemas de conocimiento 	<p>Filho, 1973; Díez, 1980; De la Cruz, 1995; Cuetos, 1990; Repetto, Téllez y Beltrán, 2002; Bofarull, 2001; Alonso Tapia y Carriedo, 1996; Sánchez, 1999; Vidal-Abarca y Gilabert, 1991; Hernández y Quintero, 2001; Meyer, 1984, 1985; Johnston, 1989; Herranz, 1987; Mora, 1993; Romero y Lavigne, 2006</p>
<p><i>Sintaxis y relaciones entre proposiciones y su interpretación</i></p>	<p>Defior y Ortuzar, 1993; Alonso Tapia y Carriedo, 1996; Cuetos, Rodríguez y Ruano, 1996; Ramos y Cuetos, 2000</p>
<p><i>Uso de la estructura y de las claves textuales para comprender</i></p> <ul style="list-style-type: none"> • Representación proposicional, relación entre proposiciones, expresiones anafóricas, etc. • Detección de ideas principales, secundarias y relaciones jerárquicas entre ellas • Conocimiento sobre tipologías textuales y cómo usa ese conocimiento para elaborar una representación del texto • Conocimiento, identificación y utilización de las claves textuales • Utilización de macrorreglas 	<p>Vidal-Abarca y Gilabert, 1991; Sánchez, 1990, 1999; Hernández y Quintero, 2001; Alonso Tapia y Carriedo, 1996; Meyer, 1984, 1985; Bofarull, 2001; Johnston, 1989; Repetto, Téllez y Beltrán, 2002; Solè, 1992; Cuetos, Rodríguez y Ruano, 1996; Ramos y Cuetos, 2000; Romero y Lavigne, 2006</p>
<p><i>Aspectos relacionados con la comprensión literal del texto</i></p> <ul style="list-style-type: none"> • Reconocimiento de ideas principales • Reconocimiento de detalles • Reconocimiento de secuencias • Reconocimiento comparativo • Reconocimiento de la causa y el efecto de las relaciones • Reconocimiento de rasgos de personalidad 	<p>Català, Català, Molina y Monclús, 2001; Cuetos, Rodríguez y Ruano, 1996; Ramos y Cuetos, 2000; Romero y Lavigne, 2006</p>

<p><i>Identificación de información relevante y capacidad para resumir lo importante</i></p>	<p>Alonso Tapia, Carriedo y González, 1992; Alonso Tapia y Carriedo, 1996; Johnston, 1989; Solè, 1992; Bofarull, 2001; Repetto, Téllez y Beltrán, 2002</p>
<p><i>Realización de inferencias/Comprensión inferencial o interpretativa</i></p> <ul style="list-style-type: none"> • Imágenes mentales que evoca • Emociones que el texto despierta • Deducción de ideas principales no explícitas • Deducción de una secuencia • Deducción de detalles de apoyo • Deducción de rasgos de personalidad • Interpretación del lenguaje figurativo 	<p>Català, Català, Molina y Monclús, 2001; Repetto, Téllez y Beltrán, 2002; Johnston, 1989; Alonso Tapia, Carriedo y Mateos, 1992; Bofarull, 2001; Hernández y Quintero, 2001; Alonso Tapia y Carriedo, 1996; Sánchez, 1999; Vidal-Abarca y Gilabert, 1991; Solè, 1992; Cuetos, Rodríguez y Ruano, 1996; Ramos y Cuetos, 2000; Romero y Lavigne, 2006</p>
<p><i>Comprensión/valoración crítica o profunda del texto</i></p>	<p>Català, Català, Molina y Monclús, 2001; Bofarull, 2001; Alonso Tapia y Gutiérrez, 1992</p>
<p><i>Percepción de la intencionalidad del autor</i></p>	<p>Johnston, 1989; Repetto, Téllez y Beltrán, 2002; Català, Català, Molina y Monclús, 2001; Alonso Tapia y Carriedo, 1996</p>
<p><i>Formulación de hipótesis/predicciones y verificación</i></p>	<p>Solè, 1992; Bofarull, 2001; Herranz, 1987</p>
<p><i>Grado en que el alumno puede manejar las fuentes escritas</i></p>	<p>Solè, 1992</p>
<p><i>El proceso de construcción del significado</i></p>	<p>Solè, 1992</p>
<p><i>Uso del contexto para comprender</i></p> <ul style="list-style-type: none"> • Contexto discursivo • Contexto de colaboración de profesores y compañeros • Reconocimiento de situaciones sociales y su influencia en el proceso 	<p>Solè, 1992; Solè y Teberosky, 2001; Moreno y García, 2000; Suárez y Meara, 1992; Edwards y Mercer, 1988; Mateos, 1991; Alonso Tapia y Carriedo, 1996; Alonso Tapia, Carriedo y Mateos, 1992</p>
<p><i>Supervisión y control del proceso de comprensión</i></p> <ul style="list-style-type: none"> • Detección de errores en la comprensión • Estrategias que utiliza para corregir sus errores 	<p>Alonso Tapia, Carriedo y Mateos, 1992; Alonso Tapia y Carriedo, 1996; Bofarull, 2001; Garner, 1987; Johnston, 1989; Solè, 1992; Moreno y García, 2000; Romero y Lavigne, 2006</p>

<p><i>El producto de la comprensión lectora: representación que se construye tras la lectura de un texto</i></p> <ul style="list-style-type: none"> • Representación textual básica • Representación situacional • Integración de conocimientos 	<p>Filho, 1973; Díez, 1980; Lázaro Martínez, 1980; Toro y Cervera, 1984; González Portal, 1985; Suárez y Meara, 1992; Herranz, 1987; De la Cruz, 1995; Johnston, 1989; Alonso Tapia, Carriedo y Mateos, 1992; Repetto, Téllez y Beltrán, 2002; Català, Català, Molina y Monclús, 2001; Cuetos, Rodríguez y Ruano, 1996; Ramos y Cuetos, 2000</p>
--	--

1.3 La metacognición y las dificultades de aprendizaje en la evaluación de la comprensión

La relevancia para el aprendizaje y el desarrollo cognitivo y socio-personal, de los procesos que permiten realizar una «mirada mental» hacia nosotros mismos, es decir, aquellos procesos que ponemos en práctica cuando pensamos sobre nuestros propios pensamientos o sobre cómo pensamos y actuamos, o cuando lo hacemos sobre nuestras propias habilidades o sobre cómo sentimos, o nos sentimos, en una situación determinada; así como de aquellos procesos que permiten realizar un control y una regulación sobre nuestra propia actividad cognitiva, ha sido considerada en grado creciente en las últimas décadas (Pozo, 1996; Mateos, 2001). Esto es así, entre otras razones, debido a los nuevos escenarios de enseñanza y aprendizaje que se han ido conformando a la luz de las transformaciones sociales acaecidas en un periodo de tiempo relativamente corto. En efecto, algunas de estas transformaciones han tenido que ver precisamente con la relatividad que ha tomado la adquisición de ciertos conocimientos, debido a la naturaleza dinámica y cambiante del entorno social y cultural, con relación a la adquisición de estrategias y actitudes que potencien precisamente la capacidad de aprendizaje en el alumnado, es decir, estrategias y actitudes que favorezcan el «aprender a aprender» (Nisbett y Shucksmith, 1987; Pozo, 1996; Pozo y Postigo, 1993).

En este sentido, el conocimiento consciente que podemos llegar a tener sobre nuestras propias estrategias para acceder a la comprensión de un texto escrito, así como el ser progresivamente capaz de manejar y controlar el uso de este conocimiento, constituye uno de los elementos fundamentales del proceso de lectura y comprensión, al tiempo que se revela clave en el proceso de su evaluación. (Alonso Tapia, Carriedo y Mateos, 1992; Alonso Tapia y Carriedo, 1996; Moreno y García, 2000; Repetto, 1997; Solè y Teberosky, 2001). Por tanto, la facilitación de su aprendizaje, así como su evaluación, se han convertido en algunos de los retos más importantes para el profesorado, ya que, por una parte, existe un amplio consenso en la comunidad científica acerca de la relación existente entre un bajo nivel de funcionamiento metacognitivo y la presencia de dificultades en el aprendizaje en general (Vygotski, 1934/1995; Nisbett y Shucksmith, 1987; Campione, Brown y Ferrara, 1987; Campione, 1987; Brown, 1980; Whitman, 1990; Mora, 1998; Swanson, 1993; Swanson y Trahan, 1996; Male, 1996; Aguilera y Rodríguez, 2003; Romero y Lavigne, 2005) y, por otra, algunos estudios han puesto también de manifiesto que bajos niveles de comprensión lectora se relacionan igualmente con la existencia de problemas en los procesos metacognitivos (Jiménez y Ortiz, 1998; Bransford, Vye, Adams y Perfetto, 1989; Alonso Tapia y Cols., 1992; Montanero, 2000).

Metacognición y Dificultades en el aprendizaje

Un pequeño recorrido evolutivo a través de los primeros años del desarrollo nos bastaría para tomar conciencia de la importancia de los procesos sociales de intercambio comunicativo en el desarrollo de la metacognición. En efecto, el desarrollo de los niños y niñas tiene lugar en contextos de interacción: interacción con aquellos que son sus progenitores, sus cuidadores o sus iguales, e interacción con los objetos físicos que tienen a su alrededor. Estos contextos de interacción provocan la progresiva apropiación, por parte del niño, de los elementos culturales, incluidas las funciones psicológicas superiores del pensamiento infantil (Vygotski, 1931/1997, 1934/1995; Rogoff, 1993). Esta apropiación no tiene lugar de manera uniforme por parte de todos los niños, sino que influyen decisivamente variables como la dotación biológica, las condiciones sociales y las experiencias de aprendizaje mediado (Feuerstein, 1991). Los adultos pueden llegar a estructurar el ambiente educativo en el que se desarrolla un niño, de tal forma, que el desarrollo de sus capacidades se vea optimizado. Una de esas formas de estructurar el ambiente educativo consiste en facilitar experiencias de interacción en la *zona de desarrollo próximo o potencial (ZDP)* de los niños. Parece ser que mayores niveles de desarrollo cognitivo en los niños, son deudores de la actitud de sus madres o padres de situarse espontáneamente en esta zona de desarrollo próximo, ayudando al niño a emitir conductas que por sí solo no podría dar (Sigel, 1982; González y Palacios, 1990; Tulviste, 1991).

Pero, ¿existen diferencias cualitativas en la adquisición de este tipo de pensamiento en los sujetos con dificultades de aprendizaje y en aquellos con necesidades educativas especiales permanentes? Se ha descrito que el proceso gradual en la adquisición del pensamiento estratégico y, en particular, del pensamiento metacognitivo, parece tener lugar en modo casi espontáneo en los sujetos sin ningún tipo de discapacidad ni problema de aprendizaje (Campione, Brown y Ferrara, 1987). Esto, obviamente, no quiere decir que tenga lugar de forma independiente a la mediación semiótica ejercida por los miembros adultos de la sociedad y la cultura de referencia, pero sí tiene que ver con que se da este desarrollo, con mayor o menor calidad, sin que exista una planificación intencional del sistema familiar ni —en muchos casos— del sistema educativo. Las situaciones de interacción, de juego, las situaciones cotidianas, el aprendizaje a través de los iguales y de los adultos de referencia, etc., pueden crear condiciones favorables para la aparición y desarrollo de conductas cada vez más estratégicas y reflexivas (Vygotski, 1931/1997; Feuerstein, 1991; Mora, 1998). Este hecho no ocurre de igual manera en los sujetos con problemas de aprendizaje y con deficiencia intelectual. En estos casos, aunque, lógicamente, con una enorme variabilidad individual, se dan determinadas circunstancias que pueden hacer peligrar la consecución adecuada de procesos de control y regulación de conductas y pensamientos y una conciencia adecuada de los propios procesos mentales. Entre estas circunstancias pueden estar aquellas ligadas a problemas en la adquisición y desarrollo de la simbolización en general y del lenguaje en particular (Vygotski, 1934/1995; Martín y Marchesi, 1990; Mora, 1998; Kurtz y Borkowski, 1987). Según Vygotski, el lenguaje se convertía progresivamente en un instrumento de control y regulación de la propia conducta, por lo que cabe esperar que niños que tienen problemas en el acceso al lenguaje y, como consecuencia, problemas en la comunicación con los otros, tengan problemas en la utilización de este instrumento para regular el comportamiento de los demás y el suyo propio. Las interacciones, a las que antes aludíamos, en las cuales los adultos guían y ayudan a los niños en la resolución de tareas, modelando de forma

espontánea o intencionada la utilización de estrategias que poco a poco serán interiorizadas y utilizadas a su vez por éstos, suponen la utilización continua del lenguaje y la comunicación (Coll y Rivière, 1985; Feuerstein, 1991). Estos aspectos, lenguaje y comunicación, son cruciales desde esta perspectiva metacognitiva en las dificultades de aprendizaje. Los problemas en la comunicación temprana, en la adquisición de un código lingüístico que progresivamente sea internalizado, repercutirán a su vez en los procesos de interiorización de las funciones metacognitivas de planificación y regulación de la propia actividad (Vygotski, 1931/1997; Martín y Marchesi, 1990; Mora, 1998).

La cuestión que se plantea, y que a todas luces es relevante en nuestro trabajo, es si conviene planificar la acción educativa para los niños con necesidades educativas especiales priorizando el aprendizaje de estrategias metacognitivas, y, como consecuencia lógica, si conviene centrar el proceso de evaluación en la actividad metacognitiva del alumnado que presenta dificultades. Numerosos estudios, algunos de los cuales quedan recogidos en el volumen de anexos (Anexo XV), han abordado este tema, sobre todo desde el punto de vista de los beneficios obtenidos especialmente por los sujetos con dificultad cognitiva cuando son trabajados preferentemente los procesos metacognitivos; pero quizás fue Vygotski quien trató este tema por primera vez, otorgando una explicación psico-evolutiva al hecho de la conveniencia de actuar, especialmente en el caso de los niños con deficiencia intelectual, sobre las funciones superiores. Al parecer, comenta Vygotski, en la mayoría de disfunciones cognitivas existen causas de orden primario, íntimamente ligadas a la propia disfunción y que, como consecuencia, son difíciles —o en algunos casos imposible— de subsanar. Tal es el caso de lesiones neurológicas, que conllevan alteraciones en la conducta, tornándola sumamente impulsiva y falta de control, o alteraciones en los procesos de memoria ligadas al retraso mental. En estos casos, las causas a las que alude Vygotski, y que serían difícilmente modificables, son la propia lesión neurológica y la propia deficiencia intelectual. Sin embargo, sus investigaciones —y posteriormente algunas llevadas a cabo por Luria (1995) en el campo de la neuropsicología— revelaron que precisamente la actuación sobre procesos superiores, que no se encontraban directamente relacionados con el «foco» de las deficiencias, es decir, con las causas de orden primario, podían mejorar significativamente las conductas y la forma de afrontar las tareas por parte de estos sujetos. De esta forma, mediante el uso del lenguaje en su función de regulación y control de la propia conducta, las funciones superiores podían actuar a modo de «prótesis de significado» y compensar las dificultades manifiestas. Vygotski escribe en 1931 (1931/1997; páginas 221 y 222):

"...las mayores posibilidades de desarrollo del niño anormal se encuentran más bien en el campo de las funciones superiores, que en el área de las inferiores. "

" La dialéctica del desarrollo y de la educación del niño anormal consiste entre otras cosas, en que no se realizan por vía directa sino indirecta. Como ya se ha dicho, las funciones psíquicas surgidas en el proceso de desarrollo histórico de la humanidad y cuya estructuración depende de la conducta colectividad del niño constituyen el campo que admite en mayor medida la nivelación y atenuación de las consecuencias del defecto y presenta mayores posibilidades para una influencia educativa. "

Por otra parte, como ya hemos indicado, los estudios sobre metacognición destacan una doble dimensión (Brown, Bransford, Ferrara y Campione, 1983): el conocimiento consciente de los propios procesos de pensamiento, incluyendo el conocimiento de las estrategias que podemos usar en una situación determinada, lo que ha recibido el

nombre de conocimiento metacognitivo o metaconocimiento; y la autorregulación y control de nuestras conductas y pensamientos. En esta segunda dimensión se incluyen habitualmente los procesos de planificación de actividades cognitivas o de aprendizaje, la supervisión de tales acciones durante su realización y la evaluación de todo el proceso y de los resultados obtenidos. Pues bien, aunque estas dos dimensiones del concepto de metacognición se hallan estrechamente interrelacionadas, parece ser que existen diferencias en cuanto a las dificultades que entraña su adquisición y en cuanto al momento evolutivo en que se consiguen de una forma más o menos eficaz (Martín y Marchesi, 1990; Campione, Brown y Ferrara, 1987; Mateos, 2001). Al parecer, la adquisición de los metaconocimientos puede tener lugar posteriormente a la de los procesos de autorregulación. Esto explicaría en parte por qué niños pequeños se muestran capaces de poner en práctica —aun de forma incipiente— habilidades de planificación y supervisión de actividades cognitivas, mientras que su capacidad de reflexión sobre esas mismas capacidades y sobre los procesos cognitivos puestos en juego sea aún muy reducida. Parece ser que tomar conciencia sobre aquello que hacemos o sobre lo que pensamos, incluidos nuestros propios procesos de conocimiento y aprendizaje, es un paso que, habitualmente, damos una vez que utilizamos esas estrategias o esos conocimientos en actividades de aprendizaje. No obstante, la puesta en práctica de procesos de autorregulación en actividades complejas como la lectura, podría implicar en mayor o menor medida, un acceso o conocimiento consciente a las propias estrategias.

En cuanto a los niños con dificultades en su aprendizaje y los niños con deficiencia intelectual, lo que acabamos de explicitar parece cumplirse de una forma más acentuada, máxime cuando, como indican autores como Vygotski, (1931/1997), Wong (1986), Nisbett y Shuksmith (1987), Martín y Marchesi (1990), Campione, Brown y Ferrara (1987), Rivière (1997), Mora (1998) o Martín (1999), las estrategias metacognitivas, tanto los metaconocimientos, como los procesos de autorregulación, deben ser enseñadas intencionalmente, facilitando su adquisición y desarrollo mediante programas instruccionales adecuados, a los sujetos que presentan algún tipo de dificultad cognitiva. En este sentido, los programas de enriquecimiento cognitivo y otros trabajos encaminados a facilitar la adquisición y desarrollo de habilidades metacognitivas especialmente en sujetos con deficiencia intelectual o dificultades de aprendizaje, han puesto de manifiesto que estos sujetos pueden beneficiarse altamente de su práctica (Meichenbaum, 1986; Kurtz y Borkowski, 1987; Campione, Brown y Ferrara, 1987; Scrugg y Mastropieri, 1992; Mastropieri y Scrugg, 1997; Mora, 1998; Saldaña, 2001). Otro aspecto importante relacionado con la adquisición y desarrollo de estas habilidades metacognitivas en sujetos con dificultades es el referido a que estos sujetos parecen beneficiarse más de los programas principalmente encaminados a facilitar la adquisición de procesos de autorregulación (Martín y Marchesi, 1990), es decir, que se dirigen a la puesta en práctica de habilidades que favorezcan el control consciente sobre un proceso de aprendizaje determinado, mediante actividades de planificación, supervisión y autoevaluación de dicho proceso y de la propia acción. En cualquier caso, parece sensato pensar que lo ideal es integrar ambas dimensiones de la metacognición en cualquier trabajo encaminado a potenciar el desarrollo de cualquier proceso de aprendizaje.

Se ha hecho también referencia a que uno de los problemas más importantes con que se encuentra cualquier intervención dirigida a la facilitación de habilidades cognitivas o metacognitivas en sujetos con dificultades de aprendizaje y, en especial, con deficiencia

intelectual, lo constituye la necesidad de transferencia y generalización de los aprendizajes (Campione y Brown, 1977; Brown, Bransford, Ferrara y Campione, 1983; Das, 1985; Kurtz y Borkowski, 1987; Ashman y Conway, 1989; Mora, 1998; Whitman, 1990; Wong, 1994; Martín, 1999). En este sentido, ser capaces de elegir la estrategia apropiada y adaptarla al momento y la situación necesarios, forma parte importante de la definición de un buen aprendizaje (Nisbett y Shucksmith, 1987; Pozo, 1996). Con relación a esta idea, muchas voces han destacado la importancia de la *adaptabilidad* como objetivo educativo básico junto con la capacidad de *transferir* lo aprendido en una situación concreta o específica, a otras en las que ese aprendizaje fuera necesario. Ambos conceptos relacionados: adaptabilidad y transferencia, se constituyen en pilares fundamentales para conformar propuestas que traten de introducir las estrategias de aprendizaje en el proceso educativo. En algunos programas de enriquecimiento cognitivo de orientación metacognitiva (Mora, 1998) y también en algunos programas de intervención psicopedagógica orientados a la mejora de las estrategias de comprensión lectora (Repetto, Sutil, Manzano, Téllez y Beltrán, 2002), se incluye la facilitación de estrategias de transferencia de los aprendizajes y la generalización de principios y leyes de carácter general entre sus objetivos, y ambos elementos están ligados estrechamente a los propios procesos metacognitivos.

Unidos a las dificultades en el desarrollo de una conciencia y un control de carácter metacognitivo, se han descrito igualmente otros problemas habituales y característicos de los sujetos que presentan algún tipo de dificultad. Entre estos se encuentran la mayor carencia de conocimientos específicos en comparación con sus compañeros (Chi, 1981; Voss, 1982; Dockrell y McShane, 1992) y los problemas de índole emocional, relacionados con el autoconcepto académico, los patrones atribucionales de éxito o fracaso, la formación de la identidad y la autoestima (De Coster, 1977; Greshman y MacMillan, 1997; Borkowski, Weyhing y Turner, 1986; Reid y Borkowski, 1987). Por su parte, Mora (1998) hace una descripción detallada de las dificultades principales que caracterizarían a la mayoría de los sujetos con dificultades de aprendizaje generalizadas, privación socio-cultural o deficiencia mental ligera. Recogemos aquí solo la enumeración de estas dificultades, ya que nos parece importante en función de referencias posteriores más relacionadas específicamente con el tema de la comprensión lectora. Estas dificultades son:

- Dificultades en el manejo de símbolos en general y en la capacidad de abstracción
- Un lenguaje caracterizado por un uso contextualizado y que no funciona habitualmente como elemento de regulación de la propia conducta, ni se utiliza en modo auto-referente
- La falta de tolerancia a la demora de reforzadores y, unido a esto, la escasa tolerancia a la frustración
- Un control externo de la conducta, lo que se relaciona con las dificultades en los procesos de autorregulación de la propia conducta
- Falta de motivación de logro
- Bajo nivel de metacognición, que, según el autor, puede estar influyendo, de una u otra forma, en los elementos restantes
- Escaso control de la impulsividad
- Una autoimagen negativa o no ajustada a la realidad
- Una inadecuada formación de la propia identidad
- Una baja capacidad de modificabilidad de sus estructuras cognitivas

Comprensión de textos y Metacognición

A lo largo de este capítulo hemos hecho referencia a un modelo interactivo como elemento explicativo de la comprensión, haciendo alusión a la interrelación existente entre los distintos procesos que intervienen en la lectura, desde procesos centrados en el alumnado a otros más textuales y contextuales; esta interacción de procesos se da también cuando un lector gestiona y controla su propio proceso de lectura y las estrategias que pone en juego para activar conocimientos previos, realizar inferencias, poner en marcha mecanismos de revisión del contenido desde lo que ya sabe, colaborar con un compañero en la comprensión y resolución de un problema o comprender una frase (Moreno y García, 2000). En efecto, desde una perspectiva metacognitiva, no cabe imaginar un modelo de lectura y comprensión que no presuponga la interacción continua de elementos y procesos implicados. Un elemento central de nuestra propuesta de evaluación-intervención es la adopción de esta perspectiva metacognitiva. En los últimos años, algunos programas de intervención en comprensión lectora han incluido entre sus propuestas la enseñanza o facilitación de estrategias metacognitivas en la lectura (Vidal-Abarca y Gilabert, 1991; Mateos, 1991; Hernández y Quintero, 2001; Repetto, Sutil, Manzano, Téllez y Beltrán, 2002); y, por otra parte, ya se ha hecho mención aquí de la relevancia que han cobrado estos procesos en la explicación de la comprensión. Aún así, en la práctica, se denota una inquietante falta de penetración en la cultura de enseñanza y aprendizaje de los centros, de las ideas, los presupuestos y las propuestas concretas de acción, expuestas en estos programas. Las razones parecen estar en consonancia con el análisis de Sánchez, Rosales y Suárez (1999) sobre la descripción o exposición de «estados ideales» de comprensión o procesos de enseñanza-aprendizaje para favorecerla, y la casi total ausencia de elementos que permitan establecer progresivas conexiones entre lo que se hace habitualmente en los centros y lo que podría llegar a hacerse. Esto, al margen de otros factores, nada desdeñables, relativos a la formación y conocimiento de técnicas y modelos prácticos y teóricos sobre la enseñanza o facilitación de la puesta en práctica de estrategias de comprensión, en general, y sobre la metacognición y las estrategias metacognitivas en la comprensión lectora, en particular. Lógicamente, estas pinceladas sobre algunas de las dificultades existentes actualmente entre parte del profesorado para adoptar métodos de trabajo que primen elementos como: la facilitación de la puesta en práctica de estrategias de aprendizaje; el apoyo a las dificultades basado en la cesión progresiva del control sobre los propios procesos de aprendizaje, y en la interacción social que fomenta la cooperación para solucionar problemas; la discusión metacognitiva; la valoración crítica y la conexión de contenidos de aprendizaje con los conocimientos y experiencias previas; etc.; no hacen sino revelar la necesidad de poner en marcha programas y actuaciones más contextualizadas y que se caractericen por la colaboración en el abordaje de las situaciones de enseñanza y aprendizaje.

Actualmente, parece existir un cierto consenso entre las distintas teorías del desarrollo y el aprendizaje acerca de que el pensamiento supone la puesta en práctica de procesos ejecutivos que faciliten la regulación de la propia actividad a través de la planificación, selección, seguimiento y evaluación de las estrategias (Martín, 1999). La explicación de la resolución de tareas en situaciones de aprendizaje ha introducido, además de factores como: (a) el conocimiento sobre las reglas y estrategias que son utilizadas por un sujeto; (b) el conocimiento acerca del objeto de aprendizaje; (c) el tipo de respuesta que se pide en un momento determinado; y (d) el contexto para adecuar la respuesta a las exigencias del mismo; otros factores relativos al conocimiento del propio

funcionamiento cognitivo y a la capacidad de ejercer control y regulación sobre los propios procesos, planificando acciones a realizar, supervisando la propia ejecución y evaluando los resultados obtenidos (Martín y Marchesi, 1990; Mora, 1998). La aportación de la teoría y la investigación sobre procesos metacognitivos ha puesto de manifiesto que lo específico de la inteligencia es esta capacidad de autorregulación consciente de los propios procesos de pensamiento y aprendizaje.

Como veíamos anteriormente, en sus escritos sobre la educación especial, Vygotski habla de proporcionar a los sujetos que manifiestan dificultades en el aprendizaje no asociadas a disfunciones graves relacionadas con la capacidad de establecer comunicación con su entorno, elementos que fueran capaces de sustituir una organización mental deficiente, es decir, elementos de control y regulación de las funciones psicológicas. En otras palabras Vygotski estaba hablando de proporcionar habilidades metacognitivas especialmente a los sujetos con dificultades en su aprendizaje (Vygotski, 1931/1997). En investigaciones posteriores, autores como Nisbett y Shucksmith (1987), Mora (1998), Meichenbaum (1986), Brown, Campione, Bray y Wilcox (1973), etc., han mostrado resultados de la investigación que venían a confirmar las tesis del psicólogo ruso, obteniendo mejoras en el funcionamiento cognitivo de estos sujetos cuando eran instruidos en el uso de habilidades de control metacognitivo.

Metaconocimiento y autorregulación del proceso de comprensión

Estos procesos metacognitivos, como hemos visto, se relacionan con el conocimiento, control y regulación de los propios procesos mentales, facilitando o, en ausencia de los mismos, dificultando el acceso de los sujetos a la comprensión del lenguaje escrito. La autoevaluación, los procesos de control y supervisión de la propia acción de la lectura, la continua creación de metas, el conocimiento sobre nuestras propias actitudes y capacidades y de cómo optimizarlas para mejorar, el conocimiento sobre el propio lenguaje, etc., parecen ser elementos muy relevantes a la hora de manejar de forma autónoma diversos tipos de textos, así como de facilitar el acceso a niveles profundos de comprensión de lo escrito. Pero ¿Cuáles son y cómo se han agrupado estas habilidades o actividades metacognitivas que nos proponemos evaluar con nuestro trabajo? Por una parte están los *metaconocimientos*, que hacen referencia al conocimiento consciente que tenemos de nuestros propios procesos mentales, de cómo pensamos, procesamos la información o de aquello que sabemos. En lo que nos concierne, los metaconocimientos son relevantes en el estudio y la evaluación de los procesos de lectura (Sánchez, Orrantía y Rosales, 1992; Van Dijk, 1980; Alonso Tapia y Cols., 1992; Vidal-Abarca y Gilabeert, 1991; Moreno y García, 2000; Montanero, 2001; Repetto, Téllez y Beltrán, 2002). Podríamos distinguir tres tipos de metaconocimiento relacionados con la lectura: (a) aquel referido al significado, sentido, objetivos o propósito de la lectura, es decir, para qué sirve leer, qué significa comprender un texto, etc. En este sentido, el conocimiento sobre los propios objetivos y propósitos con los que los alumnos afrontan una determinada tarea constituyen un elemento explicativo del éxito en esa tarea. Ya hemos aludido en este trabajo a la necesidad de que la evaluación de la comprensión lectora contenga aspectos relativos a los conocimientos sobre el lenguaje escrito y sobre la naturaleza de la lectura (Alonso Tapia y Cols., 1992). Igualmente, otros autores hacen referencia a los propósitos que guían la lectura como indicadores que discriminan a lectores estratégicos, que no presentan dificultades, de aquellos que manifiestan problemas con la lectura (Solè,

1992; Moreno y García, 2000; Solè y Teberosky, 2001; Repetto, Téllez y Beltrán, 2002); (b) otro tipo de metaconocimiento es el que se relaciona con el contenido de la lectura, es decir, la conciencia de lo que sabemos y no sabemos acerca del contenido de una lectura determinada. Por último, estaría (c) el metaconocimiento referido a los procedimientos y estrategias que podemos utilizar para acceder a la comprensión de un texto o a la resolución de una tarea propuesta.

Por otra parte, se encuentran los *procesos de autorregulación*, que son aquellos que permiten controlar y regular el propio proceso de lectura. Estos procesos juegan un papel muy importante en la comprensión, lo que se ha visto últimamente refrendado en numerosas publicaciones y programas de intervención en dificultades de lectura (Cooper, 1990; Jolivert, 1995; Sánchez, 1998, 1999; Moreno y García, 2000; Alonso Tapia y Cols, 1992; Mora, 1991; Vidal-Abarca y Gilabeert, 1991; Mateos, 1991; Hernández y Quintero, 2001; Repetto, Sutil, Manzano, Téllez y Beltrán, 2002). Básicamente se han descrito tres tipos de estos procesos: *procesos de planificación*, *procesos de supervisión* y *procesos de autoevaluación*. Ya hacíamos referencia anteriormente a que en algunos programas de enriquecimiento cognitivo se han introducido, dentro de las actividades metacognitivas, aquellas encaminadas a facilitar la *transferencia* de estrategias a contextos de aprendizaje distintos de aquel en el que se aprendieron, y la *generalización*, mediante la formulación de principios de carácter general o la integración de conocimientos en sistemas de referencia más amplios (Mora, 1998). A continuación explicamos un poco más detenidamente en qué consisten cada uno de estos procesos de autorregulación consciente de la comprensión y qué implica su puesta en práctica.

La planificación

Los procesos de *planificación* son considerados procesos metacognitivos de autorregulación, es decir, formarían parte de ese segundo ámbito de la metacognición que describieron y desarrollaron autores como Brown (1985), Campione, Brown y Ferrara (1987) o Nisbett y Shuksmith (1987). Relacionado con la comprensión de textos, un alumno pone en práctica habilidades metacognitivas de planificación cuando, previamente a su lectura, crea o selecciona unos objetivos y unas metas que guiarán su actividad, selecciona determinadas acciones que convienen a los propósitos de su lectura (por ejemplo, piensa anotar en una hoja aparte aquello más relevante del texto que va a leer), secuencia operaciones que va a realizar (por ejemplo, primero hará una lectura más o menos rápida para, posteriormente, leer de nuevo deteniéndose en aquellos aspectos más relevantes) o anticipa posibles dificultades que va a encontrar y piensa en algunas estrategias para poder solucionarlas. Tal y como apuntábamos anteriormente, estas actividades complejas implicarían, por una parte, algún grado de conocimiento consciente de sus propias estrategias y, por otra, la posibilidad de regular y controlar esas estrategias.

Como decíamos, uno de los elementos básicos del proceso de planificación consiste en crear o seleccionar el propósito o las metas que guiarán nuestra lectura. Ahora bien, la selección de objetivos o metas implica el conocimiento de los motivos que guiarán nuestra actividad, así como el conocimiento de para qué nos será útil. Esto hace referencia, por una parte, a la reflexión del alumno acerca de la tarea que se propone realizar, lo cual remitiría al *qué* se va a hacer; y, por otra, al modo en que dicha tarea se va a afrontar, lo que nos hablaría del *cómo* se va a hacer. En palabras de Flavell (1993)

esta distinción que acabamos de hacer haría referencia al *conocimiento declarativo*, es decir, saber o conocer qué se va a hacer, y, por otra parte, al *conocimiento procedimental*, o saber cómo se va a hacer algo. Según autores como Schacter y Tulving (1994), que conceptualizan los sistemas de memoria relacionándolos con las descripciones de conocimiento declarativo y procedimental, éste último se caracterizaría por no ser fácilmente accesible a lo consciente, es decir, funcionaría en modo más automático que sujeto a control consciente por parte del sujeto, lo que remite directamente a características descritas acerca de los procesos ejecutivos de control y regulación cognitiva. Mientras que el primero de estos conocimientos, el declarativo, parece relacionarse sin mayores problemas conceptuales con los metaconocimientos, el *conocimiento* procedimental ha sido relacionado, aunque con mayores reservas, con los procesos ejecutivos de control y regulación de los propios procesos (Reder, 1996). En nuestra opinión, el concepto de procesos de autorregulación, que, como se ha dicho, implica habitualmente control y regulación consciente de los propios procesos (Brown, 1985; Campione, Brown y Ferrara, 1987; Mora, 1998; Martí, 2000) reduce las posibles confusiones terminológicas y conceptuales que pueden producirse al ligar conocimiento procedimental, lo que remite al ámbito del conocimiento, con acciones cognitivas encaminadas a planificar conscientemente nuestra actividad, supervisarla y evaluarla. Por otra parte, toda acción cognitiva que suponga regulación y control *consciente* de la propia cognición, supone a su vez el conocimiento de los propios procesos y estrategias de pensamiento (Rivière, 1997), así como de las circunstancias y elementos condicionantes para llevar a cabo la actividad, lo que Paris, Lipson y Wixon (1983) denominaron *conocimiento condicional*.

El control y la supervisión del proceso de lectura y comprensión

La *supervisión* de la propia comprensión, mientras tiene lugar la actividad de lectura, constituye un elemento esencial en la evaluación de la comprensión (Mateos, 1991; 2001; Alonso Tapia y Cols. 1992; Solè, 1992; Alonso Tapia y Carriedo, 1996; Sánchez, 1998, 1999; Moreno y García, 2000). Los procesos de supervisión hacen referencia al control y regulación de la propia comprensión durante la lectura. Acciones como identificar errores e incoherencias lingüísticas en los textos, darnos cuenta de cuándo no estamos comprendiendo, poner en práctica estrategias para resolver o compensar posibles errores en la comprensión, ejecutar procesos de mediación verbal que ayuden a tomar conciencia de lo que se está entendiendo, identificar elementos de la estructura textual que apoyen la comprensión, etc., se encuadran en este tipo de habilidades.

Las estrategias que pueden ponerse en marcha para tratar de comprender un texto que se nos ha resistido en una primera lectura pueden ser variadas. Entre éstas podemos destacar volver a leer el texto con la intención y el objetivo de tratar de comprender mejor, lo que, a su vez implica reforzar los niveles de atención en la actividad. También podemos considerar, después de haber intentado comprender mejor volviendo a leer, la petición de ayuda a los profesores; en este sentido, la persistencia ante las dificultades y, por otra parte, la solicitud de colaboración, han sido aspectos destacados como positivos en programas de enriquecimiento cognitivo (Mora, 1998). La petición de colaboración, cuando se realiza posteriormente a haber intentado realizar una actividad, implica un aspecto crucial, que no hay que pasar por alto, como es la puesta en práctica de la supervisión de la propia comprensión y la evaluación ejercida sobre la misma, necesaria para ser consciente de que no se ha comprendido y solicitar ayuda. Cuando un niño no comprende y no solicita ayuda, existen muchas posibilidades de que nos

encontremos ante alguien que no es consciente de no haber comprendido, lo que además de informar de sus recursos metacognitivos, dificulta de forma importante su proceso de aprendizaje. Por otra parte, cuando se ponen en práctica mecanismos de detección de problemas en la comprensión, la utilización del contexto textual para facilitar la comprensión sin distorsionar el proceso de lectura (Alonso Tapia y Cols., 1992; Navarro y Mora, 2002), supone utilizar eficazmente el proceso de supervisión.

Igualmente, la supervisión de la propia comprensión hace referencia directamente a que el alumno, mientras lee, está pensando en aquello que está leyendo. Se enfatiza el hecho de ser conscientes de aquello que realizamos, en este caso, una tarea de lectura. Si el alumno está pensando en el significado de lo que está leyendo multiplica las posibilidades de obtener información acerca del contenido del texto, de que este contenido pueda elicitar la activación de experiencias y conocimientos previos, de poder darse cuenta de cuándo no está comprendiendo, pudiendo poner en marcha estrategias para remediar este hecho; de supervisar, en definitiva, su propia comprensión.

La puesta en práctica eficaz de estas habilidades de supervisión de la propia comprensión hace referencia a la relevancia de focalizar estratégicamente la atención durante el proceso de lectura (Moreno y García, 2000) y a comprobar que se está entendiendo el texto. Por tanto, es preciso mantener la atención durante la lectura, comprobando que se está entendiendo. En este sentido, es relevante destacar que se produce necesariamente una conexión entre un aspecto y otro, es decir, mantener la atención es una actividad que se concreta, en este caso, en una estrategia de supervisión de la propia comprensión; de ahí que hablemos de focalizar *estratégicamente* la atención. El segundo de estos aspectos alude de forma indirecta a la detección de dificultades en el proceso de comprensión, lo cual llevaría, en un lector estratégico, a tratar de encontrar remedios a esas dificultades (Alonso Tapia y Cols., 1992; Solè, 1992; Peronard, Velásquez, Crespo y Viramonte, 2002; Sánchez, 1999; Mateos, 2001).

Otro de los elementos que citábamos al comienzo como acciones prototípicas de las habilidades de supervisión es la mediación verbal del proceso de comprensión. Concretamente se alude al pensamiento para uno mismo en relación a la actividad que se está realizando. Desde los ya comentados trabajos de Vygotski (1931/1997; 1934/1995) sobre el papel progresivamente creciente de los procesos de mediación verbal en la regulación y control de la actividad por parte del niño pequeño, hasta los trabajos de Meichenbaum (1986), en los que se pone de manifiesto la relevancia del lenguaje interiorizado en la construcción de estrategias metacognitivas de control y regulación de la conducta en niños con problemas de control de la impulsividad; muchos han sido los autores que han destacado el papel de la mediación ejercida por nuestro propio lenguaje en la actividad realizada. Por otra parte, cuando exponíamos en el Cuadro 1 lo que hacían mal los alumnos que manifestaban problemas en su lectura, nos encontrábamos con una serie de estrategias o conductas poco efectivas para el objetivo de la comprensión. En este sentido, preguntar siempre lo que hay que hacer, leer rápidamente para terminar lo antes posible, etc., componen, entre otras, un abanico de acciones del alumnado con un determinado perfil de respuestas caracterizadas por una dependencia excesiva del profesor y, consecuentemente, una escasa autonomía en el trabajo y un bajo nivel de persistencia ante las dificultades (Navarro y Mora, 2002). En muchas ocasiones, a este perfil subyace el objetivo primordial de terminar la tarea, no importan las condiciones. Podría entenderse que esta visión de la realización de tareas de lectura es poco compatible con una actitud reflexiva ante el aprendizaje, que es capaz

de ampliar y valorar críticamente los contenidos del texto (Solè, 1992; 2001; Alonso Tapia y Cols., 1992) y reconstruir continuamente la representación mental a la que se ha tenido acceso, en función de la continua revisión de los contenidos a la luz de nuestros propios conocimientos previos, con la utilización consecuente de procesos de supervisión de la propia comprensión.

La autoevaluación

Al igual que los procesos de planificación y supervisión, la *autoevaluación* de la propia comprensión, habitualmente puesta en práctica cuando ha finalizado una determinada tarea de lectura, o también durante la actividad, cuando realizamos una detención para valorar en qué medida hemos entendido el texto, cuando evaluamos nuestros propios conocimientos con relación al contenido del texto, cuando valoramos críticamente su contenido, o pensamos acerca de su dificultad; constituye un elemento central de los procesos de comprensión lectora (Alonso Tapia y Cols., 1992; Alonso Tapia y Carriedo, 1996; Moreno y García, 2000; Peronard, Velásquez, Crespo y Viramonte, 2002; Sánchez, 1998, 1999; Martín, 1999; Solè y Teberosky, 2001). La autoevaluación supone poner en marcha mecanismos de detección y valoración de dificultades en nuestra comprensión, de la dificultad y facilidad global de un texto, de valoración acerca de su corrección desde un punto de vista léxico, sintáctico o semántico en función de un determinado contexto social o discursivo. Supone igualmente valorar nuestras propias estrategias puestas en juego, por ejemplo, a la hora de haber afrontado el estudio de un texto, valorar en qué grado hemos comprendido y en qué grado hemos conseguido los propósitos que habíamos formulado con relación a la tarea de lectura (Moreno y García, 2000; Repetto, Téllez y Beltrán, 2002).

Por tanto, el proceso de autoevaluación hace referencia, por una parte, a la reflexión metacognitiva sobre si se han conseguido los objetivos propuestos para la lectura, realizando una conexión con las acciones del bloque de procesos de planificación, ya que supuestamente se formularon propósitos *antes de empezar a leer*. Por otra parte, se alude a la reflexión metacognitiva acerca de si se produce una valoración crítica del contenido del texto, concretamente si se piensa al terminar de leer en el acuerdo o desacuerdo con lo que cuenta la lectura. Por último, también se hace referencia a la evaluación de la propia comprensión una vez finalizada la lectura de un texto y, en su caso, a la reflexión acerca de lo que habría que hacer para entender mejor ese texto. En síntesis, y poniendo en relación estos procesos de control y regulación consciente de la propia comprensión —en este caso los procesos de autoevaluación— con el metaconocimiento sobre estos mismos procesos, se trataría de evaluar tres tipos de acciones de autorregulación consciente y de conocimiento metacognitivo: (a) autoevaluación del grado en que se cumplieron los propios objetivos de lectura, así como sobre el conocimiento acerca de las propias estrategias para realizar la actividad; (b) autoevaluación referida a la valoración crítica acerca del contenido leído, así como sobre el conocimiento de las propias estrategias para poder abordar esta valoración; y (c) autoevaluación de la propia comprensión y, como consecuencia, sobre las estrategias a poner en práctica para comprender mejor.

La evaluación y la intervención educativa encaminadas a valorar la presencia de dificultades relacionadas con los procesos metacognitivos de autoevaluación implicados en la comprensión, así como a la mejora de estos procesos, deben dirigirse, entre otros aspectos, a discriminar entre un lector guiado por un pensamiento reflexivo y

metacognitivo, que piensa y actúa una vez terminada la lectura de un texto, que pone en marcha mecanismos de revisión del contenido desde lo que ya sabe, que valora críticamente el contenido, etc.; en contraposición a un lector que «no piensa nada más» una vez terminada la lectura, o que piensa, sencillamente, que ahora «toca» empezar otra tarea (Solè, 1992; Navarro y Mora, 2002). Tratan de distinguirse de esta forma concepciones distintas de la lectura. En este caso, las acciones mencionadas en última instancia aluden a concepciones de la lectura como simple ejercicio escolar, que no trasciende más allá, que no implica «aprender» de los textos, ya que no se piensa ni se reflexiona una vez realizada una tarea de lectura.

Procesos de transferencia y generalización

La capacidad de trasladar los conocimientos y habilidades adquiridas en un determinado contexto a otro distinto, remite habitualmente al concepto de *transferencia*. Por otra parte, la capacidad de extraer principios generales de las situaciones de aprendizaje, que nos sirvan en distintas actividades y que sean aplicables con carácter «general», alude al término de *generalización*. Ambos conceptos, han sido estudiados principalmente por los autores que han dedicado sus esfuerzos a conceptualizar y desarrollar el campo de la evaluación interactiva en la zona de desarrollo próximo; así como, estrechamente relacionado con el anterior, el de las estrategias de aprendizaje y, más concretamente, el de la metacognición (Campione, 1989; Campione, Brown y Ferrara, 1987; Campione y Brown, 1987; Nisbett y Shuksmith, 1987; Mora, 1991; 1998; Resing, 2000). Ya comentábamos anteriormente que tanto la transferencia como la generalización de aprendizajes ocupan un lugar crucial en la explicación de las dificultades de aprendizaje, en paralelo o incluidas directamente dentro de los procesos metacognitivos, y que, habitualmente, en relación a la intervención, constituyen un obstáculo difícil de pasar, en el sentido de que parecen ser estos procesos aquellos que dificultan de forma importante el desarrollo y consolidación de estrategias y habilidades, aparentemente logradas en procesos de entrenamiento metacognitivo y en el transcurso de la aplicación de programas que pretenden mejorar las habilidades de pensamiento en general (Brown, Bransford, Ferrara y Campione, 1983; Wong, 1985; Loper y Murphy, 1985; Scrugg y Mastropieri, 1992; Mastropieri y Scrugg, 1997; Navarro, 2004). Campione y Brown (1977) ya hicieron referencia a que transferir y generalizar lo aprendido a nuevas situaciones era un indicador de auténtico aprendizaje; esto, unido a su identificación como posible variable explicativa de las dificultades encontradas durante años en la mejora —y en el mantenimiento de la mejora— de los procesos de aprendizaje en sujetos con dificultades de aprendizaje o con deficiencia intelectual (Campione, Brown y Ferrara, 1987), ha ocasionado que estas estrategias hayan sido incluidas y desarrolladas en nuevos programas encaminados a enriquecer cognitivamente y metacognitivamente a estos sujetos (Mora, 1998; Ashman y Conway, 1989; Whitman, 1990; Wong, 1994).

A continuación, en el cuadro 7, recogemos un listado de las actividades de carácter metacognitivo que un sujeto puede poner en marcha a la hora de comprender un texto y que, a su vez, nos han guiado en la construcción del dispositivo de evaluación dinámica e intervención que más tarde proponemos.

Cuadro 7. Indicadores para el análisis de los procesos metacognitivos implicados en la comprensión lectora

Indicadores de Metaconocimiento sobre la lectura y sus contenidos, estrategias de lectura, actitudes y motivación

- Es consciente de que leer sirve para obtener información
- Es consciente de que leer significa comprender lo que se ha leído
- Es consciente de la estructura fonológica, léxica y sintáctica del lenguaje
- Es consciente de lo que sabe sobre elementos estructurales y tipologías de textos
- Conoce las propias estrategias que puede aplicar antes, durante y después de leer
- Conoce cómo y cuándo se aplican las propias estrategias y por qué son útiles
- Es consciente de lo que sabe acerca del contenido de una lectura
- Es consciente de lo correcto/incorrecto de un mensaje, desde un punto de vista semántico, en función del contexto
- Es consciente de su propia actitud, motivación e intereses relacionados con la lectura
- Es consciente de las propias dificultades relacionadas con la lectura, de sus expectativas y autoconcepto, así como de las posibles vías de solución a adoptar

Indicadores de Procesos implicados en la autorregulación de la comprensión

Actividades de Planificación

- Dirige por sí mismo/a procesos de reflexión y solución al problema planteado
- Identifica la tarea que debe realizar y el problema planteado
- Selecciona los recursos más adecuados para realizar la tarea
- Define y crea metas
- Anticipa errores
- Selecciona los recursos más adecuados para cada error
- Define/establece prioridades
- Secuencia operaciones a realizar

Actividades de supervisión

- Realiza actividades coherentes con la meta que ha definido, manteniendo los objetivos
- Manifiesta constancia de estrategias y de la secuencia de operaciones
- Supera obstáculos de la tarea
- Manifiesta flexibilidad en el uso de planes y estrategias
- Manifiesta control de la impulsividad. Presenta latencias
- Gestiona los errores positivamente
- Destaca los absurdos y las incoherencias lingüísticas
- Utiliza estrategias de memoria
- Supervisa sin necesidad de mediación externa constante
- Utiliza la mediación verbal durante la lectura. Se cuenta a sí mismo lo que va a hacer, lo que está haciendo o lo que ha hecho, se formula preguntas durante la tarea, plantea dudas...
- Se recupera de las interrupciones y distracciones
- Distribuye la atención de forma que se centra en los contenidos relevantes
- Controla la actividad durante la lectura para determinar si entiende o no lo que está leyendo

Actividades de autoevaluación

- Evalúa logros con relación a las metas definidas
- Evalúa la adecuación de la conducta con relación a la consecución de las metas
- Evalúa la adecuación de las metas
- Evalúa la adecuación de las estrategias
- Formula conclusiones autoalusivas
- Verifica la información obtenida de la lectura
- Se formula preguntas y dudas sobre el texto leído

Transferencia y generalización

- Formula y reconoce principios o leyes de carácter general
- Aplica estrategias o conclusiones a/de otras tareas

La evaluación de la metacompreensión lectora. Algunos estudios sobre evaluación de procesos metacognitivos implicados en la comprensión

Debemos constatar que no ha sido especialmente abundante la investigación que se ha producido con relación a la influencia de los procesos metacognitivos en la comprensión lectora. En un trabajo reciente (Navarro y Mora, 2002; Navarro y Mora, 2003a) hemos realizado una revisión de algunas de estas investigaciones. Dada la relevancia de su contenido para el trabajo de investigación que ahora presentamos, y en función de la conexión que ambos estudios han mantenido, exponemos en el Anexo XV el contenido íntegro de esta revisión, que hemos actualizado añadiendo algunas referencias de nuevos trabajos. En nuestra revisión, de la que aquí recogemos tan sólo una síntesis de algunos aspectos significativos, hemos tratado de reflejar, por una parte, el “estatus” de estos procesos en las teorías de base de los estudios, así como el modo en que se ha afrontado su evaluación, por otra. También hemos analizado algunas de las actividades específicas que se han propuesto, bien de forma independiente, o bien insertadas en propuestas de listados generales de actividades de evaluación de procesos implicados en la comprensión. En relación a un primer grupo de estudios analizados, son especialmente interesantes las investigaciones sobre el conocimiento de prelectores acerca de nociones sobre el lenguaje escrito, lo cual supone un acercamiento a la reflexión sobre elementos que permiten la comunicación y que son vistos por los niños como elementos comunicativos (Ferreiro y Teberosky, 1982; Ferreiro, 1997). En este mismo sentido, aunque con objetivos distintos, se desarrollan los estudios más específicos sobre conciencia metalingüística que han proliferado en los últimos años. Algunos de estos estudios han relacionado causalmente la capacidad para reflexionar y manipular la estructura fonológica del lenguaje hablado, y en concreto, para reflexionar conscientemente sobre los segmentos fonológicos del lenguaje oral —lo que hace referencia a la conciencia fonológica (Sinclair, Jarvella y Levelt, 1978)— en edades de 5 años, con mejores niveles de comprensión de textos al finalizar el segundo curso de la Educación Primaria (Jiménez y Ortiz, 1998). Un elemento común en estas investigaciones es la relevancia otorgada al conocimiento explícito sobre el propio lenguaje y sobre su uso en situaciones de escritura y lectura, lo que hace referencia de forma general a esta conciencia metalingüística.

Un segundo grupo de estudios lo constituyen aquellos que en su propuesta de evaluación de la comprensión incluyen la valoración de procesos metacognitivos (Gilabert y Vidal-Abarca, 1990; Hernández y Quintero, 2001; Pascual y Goikoetxea, 2003). Muchas de estas investigaciones se han dirigido especialmente a valorar el uso y la conciencia que tenían los alumnos acerca del conocimiento sobre las estructuras textuales (Richgels y otros, 1987; Cook y Mayer, 1988; Scardamalia y Bereiter, 1984). En general, estos estudios muestran la relevancia para la comprensión del conocimiento sobre la estructura textual y de su adecuada identificación, facilitando, entre otros aspectos, la realización de inferencias. Pero hay algo más, parece aun más importante y explicativo de la comprensión, la conciencia que los niños desarrollan sobre los esquemas estructurales, por ejemplo sobre los esquemas de las historias.

Entre las actividades específicas que se han utilizado para evaluar estas habilidades metacognitivas implicadas en el proceso lector se encuentra la detección de errores en un texto (August, Flavell y Clift, 1984; Garner, 1987), provocando que se pongan en marcha mecanismos de supervisión para dar cuenta de los errores introducidos (léxicos, sintácticos o semánticos) y para dar solución a los mismos. En otras pruebas planteadas,

los sujetos deben comparar textos en función de la dificultad de su comprensión, provocando un análisis sobre dónde se encontraban estas dificultades y sobre cómo solventarlas. También puede pedirse al alumnado que expliciten verbalmente las estrategias que deben ponerse en marcha para comprender un texto. En ocasiones se pide a un sujeto que ayude a otro compañero en una tarea de lectura, esperando que verbalice estrategias de comprensión para apoyar la actividad (Garner, Macready y Wagoner, 1984). Por otra parte, también se recoge información sobre la puesta en práctica de estas habilidades mediante la observación sistemática y la entrevista metacognitiva durante la realización de las diversas tareas propuestas (Olshavsky, 1976; Garner y Alexander, 1982; Hare y Smith, 1982). En este sentido, se trata de aprovechar los momentos de resolución de las tareas por parte del alumnado para realizar preguntas que nos orienten acerca del conocimiento que tiene sobre su proceso de comprensión y sobre las estrategias que conoce, y que nos den información sobre la puesta en práctica de procesos de autorregulación.

Hay un tercer grupo de estudios revisados en el que, en función de la teoría de base con la que se afrontan los trabajos, se contemplan los procesos metacognitivos explícitamente, como parte fundamental a evaluar en la comprensión de textos. (Alonso Tapia, Carriedo y Mateos, 1992; Alonso Tapia, Carriedo y González, 1992; Montanero, 2000; Montanero, Blázquez y León, 2002; Peronard, Crespo y Velásquez, 2000; Peronard, Crespo y Guerrero, 2001; Téllez, 2004). De muchos de estos estudios podría desprenderse que la capacidad para controlar y regular el propio proceso de comprensión, utilizando adecuadamente estrategias eficaces, se relaciona positivamente con una mayor comprensión de los textos. Para que este uso hábil de las propias estrategias tenga lugar, se hace necesario, igualmente, el conocimiento por parte de los alumnos acerca de sus propias estrategias, de cuándo utilizarlas y del por qué son útiles. Ambas dimensiones de la metacognición parecen estar íntimamente relacionadas. De la misma forma, algunos autores, en sus estudios, hacen referencia a situaciones de colaboración en la resolución de las tareas planteadas, así como a la necesidad de contextos de interacción y mediación del profesor en las actividades propuestas para que tengan lugar las mejoras pretendidas (Montanero, 2000). En el siguiente capítulo nos adentraremos de lleno en los diversos modelos y propuestas de evaluación dinámica, que han destacado especialmente este componente de *interacción mediada*, que facilitaría la obtención de información relevante y útil sobre las habilidades, dificultades y necesidades educativas de los sujetos, al tiempo que permite deducir el tipo de ayuda más ajustada a estas necesidades.

2. Evaluación dinámica

2.1 Contextos de desarrollo potencial y conceptos clave en la evaluación dinámica

Introducción

La evaluación dinámica (ED) surge en el contexto de los estudios del psicólogo ruso Lev Vygotski sobre el desarrollo de conceptos en los niños y la apropiación de los instrumentos simbólicos propios de la cultura de referencia (Vygotski, 1934/1995). Vygotski constata que se encuentran diferencias significativas entre la capacidad mostrada por diferentes niños, con igual capacidad intelectual medida con pruebas de CI, a la hora de beneficiarse de la guía y ayuda de un adulto o compañero más experto. Las diferencias entre el rendimiento del alumno sin la guía y ayuda del experto y con ésta, constituirían una *zona de desarrollo próximo o potencial (ZDP)*. Una de las interpretaciones cualitativas más interesantes de la ZDP, y con relevantes implicaciones educativas, es la de que constituye el reflejo de la capacidad del niño para beneficiarse del apoyo ofrecido por el adulto y también del aprendizaje en colaboración (Brown y Ferrara, 1985; Campione y Brown, 1987; Campione, 1989; Rogoff, 1993). Las diferencias encontradas, según Vygotski, podían influir decisivamente en el aprovechamiento y rendimiento posterior en el contexto escolar, por lo que la medida ofrecida por el análisis de las capacidades que los alumnos podían ejecutar en contextos de interacción mediada, sería de mayor utilidad que las medidas estáticas de edad mental. Todos los métodos y propuestas de ED que han sido desarrollados desde entonces tienen en común la inclusión de actividad mediadora por parte del sujeto evaluador, con el fin de obtener el máximo de información valiosa acerca de los procesos estudiados (Feuerstein, Rand y Hoffman, 1979; Lidz, 1987; Haywood y Tzurriel, 1992).

Algunas de las influencias más significativas que tuvieron los estudios y escritos de Vygotski en los procesos de enseñanza y aprendizaje tienen que ver con una nueva consideración de las situaciones de aprendizaje, de los agentes de aprendizaje y de las relaciones que mantendrían ambos elementos (Wertsch, Tulviste y Hagstrom, 1993; Cubero y Luque, 2001). La consideración de situaciones diádicas o grupales como agentes de aprendizaje ha tenido importantes implicaciones prácticas, tanto para la intervención educativa como para la evaluación. Por una parte, se han identificado formas y situaciones de aprendizaje compartido y en colaboración (Rogoff, 1993). En el aprendizaje en colaboración, los alumnos no son destinatarios pasivos de unos conocimientos ofrecidos por el profesor, ni tampoco son aprendices independientes que son capaces de llegar a sus propias soluciones; en este caso se convierten en actores de situaciones de aprendizaje compartidas por ellos mismos y también por los profesores (Rogoff y Wertsch, 1984; Newman, Griffin y Cole, 1989). Por otra parte, se ha demostrado que la evaluación de la competencia de los alumnos proporciona una mayor y más rica información en situaciones en las que éstos participan en actividades de aprendizaje asistido (Brown y Ferrara, 1985). Estas actividades proporcionarían el mejor modo de acercarse a los procesos aún no manifiestos en el alumno, pero que es

capaz de actualizar con la ayuda ajustada. Así mismo, Vygotski sostenía que los psicólogos no debían limitarse a efectuar un análisis de las funciones que ya formaban parte del repertorio cognitivo del sujeto, es decir, de aquellas funciones que ya habían madurado; debían considerar igualmente aquellas que se encuentran en proceso de ser incorporadas por el sujeto a su bagaje personal. En otras palabras, Vygotski argumentaba que no solo había que tener en cuenta el nivel de desarrollo actual del sujeto, sino que también había que considerar la zona de desarrollo potencial para acceder al nivel de desarrollo potencial (Vygotski, 1934/1995). Las funciones cognitivas en proceso de desarrollo podían ser evaluadas mediante la realización de tareas en las cuales el niño tuviera que colaborar con adultos o compañeros más aventajados para poder solucionarlas; estas situaciones servirían de base para evaluar la diferencia entre lo que el niño era capaz de hacer sin ayuda (su nivel de desarrollo actual) y lo que se mostraba capaz de hacer con la ayuda de otros (su nivel de desarrollo potencial).

La zona de desarrollo potencial puede ser interpretada como una *región dinámica sensible al aprendizaje de aptitudes culturales* (Newman y Holzman, 1993; Rogoff, 1993; Wertsch, 1988). Es a través de esta ZDP donde tendrían lugar los mecanismos de transmisión del conocimiento cultural desde los miembros expertos de la cultura a los miembros que se incorporan a ésta (Cole, 1985). Otra interpretación posible de la ZDP, según Sternberg y Grigorenko (2003) estaría más relacionada con los procesos educativos escolares. Desde esta concepción se alude a la relevancia de tomar en consideración los niveles de desarrollo actuales del alumnado para extender progresivamente zonas de desarrollo próximo o potencial en la medida que los alumnos van respondiendo positivamente a la intervención planteada (Wertsch, 1991) y, por tanto, van incorporando a su propio repertorio los contenidos trabajados. Se trata, como proponen Cole (1985), Nisbett y Shucksmith (1987), Mora (1998), o Coll y Solè (2001), entre otros, de promover una progresiva asunción del control de la actividad.

Aún una interpretación más de la ZDP, que amplía el concepto y nos parece especialmente interesante dadas las características de nuestro estudio. Se trata de la propuesta que formularon Newman, Griffin y Cole (1989), en la que formularon la noción de *zona de construcción* para referirse al espacio compartido entre dos personas con desigual experiencia que tratan de alcanzar un objetivo común, y en el que pueden tener lugar procesos interpsicológicos. Aunque la explicación de cómo se produce el aprendizaje no difiere de la ofrecida para la ZDP, el concepto remite directamente a situaciones de pensamiento en interacción que han supuesto una línea de trabajo e investigación de gran interés para la psicología educativa en los últimos años (p.e. Rogoff, 1993; Aguilera, 1997). Mencionábamos anteriormente que en los trabajos de Vygotski podíamos encontrar su propuesta de evaluación de las capacidades en desarrollo mediante la intervención en la ZDP. Probablemente ha sido este concepto, como medio individualizado para optimizar los procesos de evaluación, el que ha sugerido la mayoría de los procedimientos y las pruebas de evaluación dinámica.

Interacción mediada como concepto clave en la evaluación dinámica

Los instrumentos psicológicos y la mediación semiótica

Según Vygotski, el desarrollo de las funciones psicológicas superiores exigía la apropiación e interiorización de instrumentos psicológicos y signos en contextos de interacción (Vygotski, 1934/1995). Esta afirmación supone tanto como decir que esto

que se produce en contextos de interacción es *aprendizaje* y que éste —bajo circunstancias relativas a su vez al grado de maduración biológica del sujeto— genera desarrollo (Rivière, 1985). El desarrollo tendría lugar por tanto en la interacción que mantienen sujetos más expertos culturalmente con aquellos que se incorporan a una determinada comunidad. Los sujetos más expertos median entre los estímulos, ya sean materiales o sociales, y los sujetos menos expertos. Así tiene lugar un proceso de construcción de los contenidos culturales generados en la vida social y, a su vez, contruidos y acumulados a través de la historia (Vygotski, 1934/1995, 1979/1996; Wertsch, 1988; Feuerstein, Rand y Hoffman, 1979; Feuerstein, 1991; Cubero y Luque, 2001). En relación con lo anterior, se ha mostrado en numerosas ocasiones que la ausencia de mediación generadora de desarrollo —lo que en palabras de Feuerstein constituye la *experiencia de aprendizaje mediado*—, es decir, ausencia o pobreza de estimulación cognitiva/metacognitiva, lingüística y socio-afectiva por parte de los cuidadores principales de un sujeto, o ausencia de estos mismos elementos en situaciones de aprendizaje escolar, puede estar en la base de diversos tipos de problemas de aprendizaje y situaciones de deprivación socio-cultural (Sewell y Price, 1991; Feuerstein, 1991; Mora, 1998; Kozulin, 2000). A este respecto, Kozulin plantea que la investigación desarrollada mediante la utilización de pruebas dinámicas, con su característica fundamental de establecer una mediación en el proceso de evaluación, puede contribuir eficazmente a establecer un diagnóstico diferencial de las causas que provocan las dificultades en el aprendizaje. La mediación intensiva que suele proponerse en los modelos dinámicos puede indicar, en el caso de una respuesta positiva del sujeto en contraposición a su respuesta ante otras tareas no mediadas, que la ausencia de adecuadas experiencias de aprendizaje mediado constituirían la causa principal/original de las dificultades encontradas.

En elaboraciones más recientes, se han tratado de especificar cómo y de qué forma algunos elementos o procesos intervienen concretamente en las interacciones relacionadas con la transmisión de la cultura. Se han descrito dos procesos clave en las interacciones mediadoras de los miembros más expertos de la cultura con los menos expertos que favorecerían este proceso de construcción de los contenidos culturales: la *facilitación de la comprensión* y la *cesión del control* (Luque, Ortega y Cubero, 1997). Cuando hablamos en educación de situarse en la ZDP de un alumno o compañero menos aventajado para elaborar contenidos conjuntamente, utilizando ayudas ajustadas a su nivel de funcionamiento cognitivo o socio-moral, facilitándole el progresivo acceso autónomo a esos mismos contenidos, estamos actualizando estos procesos de forma efectiva. La evaluación de este proceso, que constituye un *continuum* desde la actividad compartida, donde el control de la actividad puede ser ejercido por el miembro más experto, hasta la actividad autónoma, donde se da una cesión del control de la actividad al miembro en principio menos experto, constituye una fuente relevante de información acerca de los procesos cognitivos o afectivos que se encuentran en maduración y a los cuales se puede acceder en el momento con la ayuda necesaria y ajustada. Se consigue así acercarse a la evaluación dinámica, y no estática, de procesos de adaptación cognitiva y afectiva que, por su naturaleza, se encuentran en continuo movimiento.

La actividad desarrollada por el evaluador en el marco de un proceso de evaluación dinámica guarda una estrecha relación con el papel que pasó a desempeñar el docente en un contexto educativo presidido por un enfoque sociocultural; éste, tal y como señala Kozulin (2000), en lugar de erigirse en una figura autoritaria que dirige a voluntad el proceso de enseñanza, pasó a convertirse en un miembro *experto* de un grupo de trabajo

e investigación en el que participaba cumpliendo funciones que se encontraban más cerca del asesoramiento y la guía. El devenir y la lógica de la resolución de un problema de aprendizaje —una tarea de lectura por ejemplo— constituirían la referencia y el carácter de la participación del educador desde esta perspectiva. Por otra parte, este papel desempeñado por el docente dentro del aula y también con respecto a las relaciones establecidas en el proceso educativo, papel caracterizado por la asunción de funciones de referencia para el grupo, guía del proceso y también de agente que sustenta en un principio el control de la actividad, puede llegar a ser interiorizado por el alumno, integrando como funciones propias las que han sido desplegadas primeramente por el docente. Según Vygotski (1979/1996), muchas de las relaciones que tienen lugar en una actividad realizada en grupo pasarían a formar parte, mediante un proceso de interiorización, de las funciones mentales individuales en forma de relaciones entre los propios procesos intelectivos. Más concretamente, Wertsch y Stone (1985) hacen alusión a que las características de tipo estructural del funcionamiento cognitivo interpersonal, como por ejemplo la organización dialéctica pregunta-respuesta, pasarían a formar parte integrante del funcionamiento cognitivo intrapersonal o intrapsicológico, mediante un proceso progresivo de interiorización mediada.

Tal y como propuso Vygotski (1934/1995, 1979/1996) el desarrollo de los procesos psicológicos superiores era el resultado de la actividad mediada. Vygotski hizo referencia principalmente a dos elementos que podían generar una función mediadora de estos procesos: los *instrumentos psicológicos* y los *seres humanos*. En cuanto a los primeros, su desarrollo tuvo lugar en función de la actividad de representación simbólica asignada a los instrumentos materiales mediante su empleo colectivo y la comunicación interpersonal (Kozulin, 2000). Estos instrumentos psicológicos tienen como objeto la mediación semiótica entre los procesos psicológicos humanos, trascendiendo el instrumento material. La escritura, la lectura, los sistemas de notación numéricos o los diversos lenguajes informáticos, serían complejos y sofisticados instrumentos psicológicos de mediación. Como expresaron Cole y Scribner (1996) en la introducción a la primera edición inglesa de *El desarrollo de los procesos psicológicos superiores*, Vygotski estaba convencido de que el uso y la interiorización de los sistemas de signos construidos históricamente, debía generar transformaciones significativas en el individuo y crear vínculos entre formas tempranas y posteriores del desarrollo humano. De esta forma, el desarrollo evolutivo se enraizaba, según el psicólogo ruso, en la historia social y cultural de la especie.

En lo relativo a la mediación de las funciones psicológicas ejercida por los seres humanos, Vygotski hizo referencia a que ésta se producía en un contexto de interacción sociocultural, donde tiene lugar una continua interpretación de determinadas acciones realizadas por los miembros que se incorporan a una comunidad por parte de los miembros adultos y expertos, *signándolas* —en el sentido de transformarlas en signos— y, por consiguiente, dotándolas de significado y transformándolas sustancialmente de esta forma. La mediación queda así ligada al concepto de *interiorización* o *internalización*, que vendría a ser la reconstrucción interna de una operación que ha tenido lugar primeramente en forma externa al sujeto. Este proceso queda expresado por el mismo Vygotski de la siguiente forma (Vygotski, 1979/1996; pág. 93):

Un proceso interpersonal queda transformado en otro intrapersonal. En el desarrollo cultural del niño, toda función aparece dos veces: primero a nivel social, y más tarde, a nivel individual; primero entre personas (interpsicológica), y después, en el interior del propio niño (intraprológica). Esto puede aplicarse igualmente a la atención voluntaria, a

la memoria lógica y a la formación de conceptos. Todas las funciones superiores se originan como relaciones entre seres humanos.

La experiencia de aprendizaje mediado

Por su parte, la *teoría de la experiencia de aprendizaje mediado* (EAM) (Feuerstein, Rand y Hoffman, 1979; Feuerstein, Rand, Hoffman y Miller, 1980; Feuerstein, Rand y Rynders, 1988; Feuerstein, 1991), desarrollada por el psicólogo israelí Reuven Feuerstein y sus colaboradores, presenta un desarrollo del concepto de mediación íntegramente centrado en la acción del ser humano. La EAM está definida por el propio autor como *una cualidad de la interacción ser humano-entorno que resulta de los cambios introducidos en esta interacción por un mediador humano que se interpone entre el organismo receptor y las fuentes del estímulo. El mediador selecciona, organiza y planifica los estímulos, variando su amplitud, frecuencia e intensidad, y los transforma en poderosos determinantes de un comportamiento, en lugar de estímulos al azar cuya aparición, registro y efectos pueden ser puramente probabilísticos* (Feuerstein, 1991, pág. 38). En efecto, la relevancia que se otorga a la mediación intencional de los procesos psicológicos desde esta teoría es tal que de ella se deriva una concepción de las diferencias cognitivas individuales basada en los diferentes grados de exposición a experiencias de aprendizaje mediado. De este modo, aspectos tales como la herencia y las condiciones del organismo más ligadas a ésta, el nivel madurativo o de configuración organizacional del organismo, el desarrollo socioemocional de los niños y de sus padres, los estímulos ambientales, el estatus socioeconómico y el nivel educativo, y las diferencias culturales, son considerados como factores *distales* en la etiología del desarrollo cognitivo y su influencia no explicaría de forma inevitable las diferencias individuales. Por el contrario, la presencia o ausencia del factor *proximal*, es decir, la EAM, explicaría un desarrollo cognitivo adecuado, acompañado de una alta capacidad de modificabilidad de las propias funciones cognitivas o, en su ausencia, un *síndrome de desarrollo cognitivo inadecuado de privación sociocultural*, así como una capacidad reducida para modificar sus propias experiencias y capacidades.

Lógicamente, determinados factores distales podrían explicar carencias en la posibilidad misma de beneficiarse de la EAM, pero la constatación de que condicionantes distales endógenos o exógenos similares muestran correlaciones muy limitadas con el desarrollo posterior cuando son tenidos en cuenta factores relativos a las EAM (Klein y Feuerstein, 1985; Klein, Weider y Greenspan, 1987) revelan la importancia de la EAM como factor decisivo en la explicación de las diferencias en capacidad de desarrollo y aprendizaje.

La experiencia de aprendizaje mediado es, como apuntábamos hace un momento, un proceso interactivo a través del cual un miembro experto de la cultura se interpone —media— entre los estímulos y el niño, introduciendo modificaciones, tanto a nivel de la tarea que se pretende mostrar —mediante la selección y organización de los estímulos, la planificación de la tarea o la frecuencia e intensidad de la misma, etc. —, como a nivel del propio niño, ya que se trata de conducirlo, a través de esta interacción mediada, a formas de pensamiento más complejas. Un análisis que nos parece muy interesante en torno a este concepto, que puede ser compartido con el concepto de mediación semiótica de Vygotski, es el que realizan Sternberg y Grigorenko (2003), confrontando la EAM con los procedimientos de instrucción directa (Bauman, 1986; Bruning, Schraw y Ronning, 2002); en ésta —sin dejar de constituir un proceso de mediación—, el adulto *enseña directamente* hechos, conceptos o procedimientos mediante la verbalización del «qué es» o del «cómo se hace», o mediante una

demostración práctica; en cambio, la enseñanza que se desprende de la EAM es una enseñanza más *indirecta*, que se realiza a través de la mediación de la experiencia, estimulando habilidades de pensamiento y razonamiento que van más allá de la tarea.

La EAM tendría como finalidad incrementar las posibilidades de un individuo para beneficiarse de la exposición *no mediada* a los estímulos. Aunque Feuerstein solo habla de crear una disposición actitudinal en el receptor de la EAM para poder beneficiarse de ulteriores exposiciones directas a los estímulos, la idea podría ampliarse en línea con el concepto anteriormente visto de interiorización de Vygotski (1979/1996), es decir, la EAM podría favorecer la interiorización de determinadas funciones que a su vez favorecerían el aprendizaje, como por ejemplo funciones de tipo metacognitivo. También podemos encontrar registros más recientes, que argumentan en torno a esta idea, cuando hacemos referencia a la cesión progresiva del control y la regulación de la actividad desde el adulto hacia el niño (Martín y Marchesi, 1990; Luque, Ortega y Cubero, 1997; Mora, 1998; Coll y Edwards, 1996; Vermunt y Verloop, 1999; Mercer, 2000; Colomina, Onrubia y Roquera, 2001; Coll y Solè, 2001).

Relacionado con esta capacidad de beneficiarse cada vez más y mejor de la exposición directa a los estímulos y también a las situaciones de aprendizaje mediado, estaría el concepto, apuntado unas líneas más arriba, de *modificabilidad estructural cognitiva* (Feuerstein, 1991, 1998a, 1998b). Se trata de un concepto que remite a su vez a una concepción dinámica del constructo inteligencia y del proceso de aprendizaje. El grado de modificabilidad estructural cognitiva de un sujeto en la teoría de Feuerstein sirve a los propósitos de establecer la naturaleza, frecuencia e intensidad de los estímulos requeridos —la EAM— para producir cambios significativos en la organización y funcionamiento cognitivo de un individuo en un área de conocimiento determinada.

Entre los criterios o características que han sido asignados a esta EAM, como artífice del desarrollo de la modificabilidad estructural cognitiva en los sujetos (cuadro 8), han sido destacados por su autor principalmente tres: la *intencionalidad*, la *trascendencia* y la *mediación del significado* (Feuerstein, 1991; Kozulin, 2000). Estos tres elementos, que tendrían una dimensión universal según Feuerstein, asegurarían la formación de esquemas de pensamiento y acción flexibles y, como consecuencia, la modificabilidad estructural cognitiva como característica específicamente humana. El primero de ellos, la *intencionalidad*, hace referencia a que el adulto que media entre el estímulo y el niño (adoptando un ejemplo de situación típicamente escolar) lo hace de forma intencionada y no incidental. Los estímulos son organizados y preparados previamente, y la situación de aprendizaje es organizada e incluso manipulada con objeto de que sirva al propósito de facilitar un determinado aprendizaje. Un aspecto relevante de esta característica lo constituye el hecho de que el objetivo básico de la interacción mediada intencionalmente no son los objetos o estímulos concretos que se están manejando, sino los procesos de pensamiento. El adulto se centra, no solo en aspectos relativos a la actividad observable, sino, sobre todo, en aspectos relativos a procesos cognitivos como los estados de atención, los procesos de resolución de un problema, los errores cometidos, etc.

Cuadro 8. Criterios para las Experiencias de Aprendizaje Mediado. Tomado de: *El modelo de Experiencia de Aprendizaje Mediado (EAM). Jerusalén, 1998*

1. Intencionalidad y reciprocidad
2. Trascendencia
3. Mediación del significado
4. Mediación del sentido de competencia
5. Regulación y control mediado de la conducta
6. Mediación del acto de compartir la conducta
7. Individualización mediada y diferenciación psicológica
8. Mediación de la búsqueda, planificación y logros de los objetivos de la conducta
9. Mediación del cambio: implica búsqueda de novedad y complejidad
10. Mediación del conocimiento del ser humano como entidad cambiante
11. Optimismo

La segunda de las características citadas, la *trascendencia*, sitúa el énfasis en que las actividades mediadas intencionalmente pueden tener un sentido —una trascendencia— mayor y más importante que el que se pone de manifiesto en apariencia. De esta forma, determinados aprendizajes que efectúan los niños en contextos familiares o escolares poseen un valor que va más allá de la situación y del momento en el que se producen (Kozulin, 2000; Navarro, 2005). Esta cualidad de la EAM no requiere control consciente y deliberado, según sus autores, para cumplir su función; de esta forma muchos niños y niñas en la escuela o en sus barrios podrían aprender el valor del trabajo en equipo cuando participan de juegos o deportes colectivos, aunque esos aprendizajes no fueran planeados ni estén controlados deliberadamente. Más tarde volveremos sobre este punto, cuando abordemos nuestra propuesta de evaluación dinámica, ya que el hecho de que las actividades mediadoras estén planteadas y organizadas de forma controlada, regulada y consciente o, por el contrario, carezcan de estas características, va a tener una importancia significativa en función de aspectos como los contenidos del aprendizaje y los sujetos del mismo.

Por su parte, la *mediación del significado* hace referencia a la necesidad de dotar de significado a la actividad mediadora (Kozulin, 2000). Una EAM sólo tendría lugar cuando los estímulos, la situación o los contenidos poseen significado para el aprendiz; de este modo, la potencia de una intervención dirigida a mediar un determinado aprendizaje se vería reducida o incluso diluida en ausencia de significatividad para quien aprende. Esta cualidad de la EAM remite inevitablemente al concepto y la teoría del *aprendizaje verbal significativo* desarrollado por Ausubel (Ausubel, 1963; Ausubel, Novak y Hanesian, 1978), aunque desde esta teoría se hace especial hincapié en la relevancia de los conocimientos previos del aprendiz y de su actitud con relación a los nuevos aprendizajes (Novak, 1998). El ajuste y la interacción entre el material presentado, los conocimientos previos y la actitud ante el aprendizaje, explicaría el aprendizaje significativo del sujeto (Martín y Solè, 2001). Sin embargo, en la revisión de Novak (1998) sobre la teoría de Ausubel, éste sí plantea la importancia de la *interacción educativa* en el proceso de enseñanza y aprendizaje. De esta forma, aunque la acción educativa eficaz desde esta teoría se basa en la exposición de la información mediante la selección, estructuración y secuenciación adecuada de los contenidos de

aprendizaje, Novak destaca el intercambio comunicativo de significados y sentimientos, y la negociación compartida de significados como fuentes de aprendizaje significativo.

Las aptitudes como formas de experiencia en desarrollo

Hasta ahora hemos analizado algunas claves relativas a los procesos de interacción mediada que fundamentan y tienen lugar en los procesos educativos escolares. Hemos revisado cómo esta interacción mediada actúa en la ZDP del sujeto con el objetivo de conducirlo hacia un nivel de funcionamiento cognitivo superior. Las habilidades o aptitudes que integran el repertorio cognitivo del sujeto pueden verse de esta forma modificadas mediante un proceso de interiorización progresiva de estructuras sociales en su origen que, en definitiva, darían cuenta de la flexibilidad y plasticidad de nuestras estructuras cognitivas y de las aptitudes que contiene. En este sentido, además de la interacción mediada y la consideración de una ZDP, otro aspecto fundamental de la evaluación dinámica es, precisamente, su concepción *dinámica* del desarrollo y el aprendizaje. A adentrarnos en este concepto y sus derivadas vamos a dedicar las páginas siguientes.

A parte de las formuladas por Vygotski y Feuerstein, una conceptualización dinámica de los procesos cognitivos y socioemocionales que se ha abierto camino en estos últimos años ha sido la formulada por el psicólogo estadounidense Robert Sternberg. En el *modelo de experiencia en desarrollo* formulado por Sternberg (1998, 1999b; Sternberg y Grigorenko, 2003), las aptitudes —incluido el constructo inteligencia— no son una entidad inamovible y prefijada de antemano, sino que se caracterizan por su plasticidad y flexibilidad (Feuerstein, 1991; Sternberg y Grigorenko, 2003). La naturaleza de las aptitudes sería dinámica y su desarrollo estaría en función de la actividad mediada (Vygotski, 1934/1995, 1979/1996). Así entendidas, las aptitudes pasarían a ser formas de experiencia en desarrollo. Sternberg define la experiencia en desarrollo como *el proceso continuo de adquisición y consolidación de un conjunto de aptitudes necesarias para un nivel elevado de experiencia y dominio en uno o más ámbitos de actuación en la vida* (Sternberg y Grigorenko, 2003, pág. 17). Las pruebas convencionales de aptitud medirían sólo una parcela limitada de esta experiencia en desarrollo.

Desde una perspectiva que considera las aptitudes como formas de experiencia en desarrollo, la evaluación de una de ellas debería dar cuenta de diferencias cualitativas en los tipos de experiencia en desarrollo surgidas; esto podría argumentarse también de acuerdo con lo que sería una evaluación de los distintos procesos implicados en un dominio específico como por ejemplo la comprensión de un texto y el distinto grado de desarrollo que puede llegar a observarse en algunos de ellos. También desde esta perspectiva, Sternberg argumenta que la evaluación de la experiencia en desarrollo debería mostrar una alta correlación con la capacidad y el éxito posterior, ya que ambas formas de experiencia compartirían la necesidad de poner en práctica lo que él mismo denominó *metacomponentes* (Sternberg, 1985), es decir, habilidades estratégicas de control y regulación del pensamiento, lo que se revela como un aspecto esencial en cualquier contexto de aprendizaje. Por otra parte, Sternberg no observa diferencias esenciales entre las *aptitudes* que han sido objeto de valoración por parte de las pruebas convencionales de inteligencia, y aquellas que conforman los componentes del logro académico o profesional. Todas ellas serían consecuencia de la covariación y la

interacción entre factores genéticos y ambientales, y todas son en última instancia formas de experiencia en continuo desarrollo.

El modelo de Sternberg plantea la interacción de seis elementos esenciales, algunos de los cuales guardan relación con su teoría triárquica de la inteligencia (Sternberg, 1985). Entre estos elementos se encuentran: (1) las *aptitudes metacognitivas*, o metacomponentes en la teoría de Sternberg; (2) las *aptitudes para el aprendizaje*, como componentes relacionados con la adquisición de conocimientos; (3) las *aptitudes para el pensamiento*, que se dividen en aptitudes para el pensamiento crítico/analítico, para el pensamiento creativo y para el pensamiento práctico; (4) el *conocimiento*, que puede ser declarativo o procedimental; (5) la *motivación*, que puede estar orientada al logro o a la competencia personal; y (6) el *contexto*. Como apuntábamos más arriba, todos estos elementos mantienen una interacción constante y total, influyéndose mutuamente en la configuración de determinados tipos de experiencia en desarrollo para ámbitos determinados. Según Sternberg, sólo una evaluación dinámica, que tenga en cuenta que las distintas aptitudes que queramos medir son en sí experiencia continua en desarrollo, y que aporte información sobre el potencial de este desarrollo, no sólo a nivel cognitivo, sino también social y personal, podría mostrarse eficaz para recoger información relevante sobre el perfil o la competencia de un sujeto en un área de conocimiento determinada o sobre su capacidad de desarrollo sociopersonal.

En definitiva, una de las extensiones más claras y acertadas de las ideas de Vygotski acerca de la necesidad de evaluar los procesos cognitivos que se encuentran en maduración, y no sólo los que ya forman parte del repertorio cognitivo del sujeto, es la que formulan Sternberg y Grigorenko (2003). Estos autores plantean que una de las razones fundamentales por las que las pruebas convencionales y estáticas de inteligencia, o de otras aptitudes, no pueden ofrecer una información adecuada de la competencia del sujeto, es debido a que, al no tener en cuenta el carácter procesual de las aptitudes, es decir, que éstas se encuentran en continuo desarrollo, terminan por mezclar la medición de aptitudes ya desarrolladas con la de aptitudes en proceso de desarrollo. No podemos saber con la aplicación de estas pruebas en *qué medida* estamos evaluando las distintas capacidades que se ponen en juego a la hora de resolver las tareas de la prueba. En opinión de estos y de otros autores, las pruebas de evaluación dinámica podrían acceder a esta información (p.e. Brown y Ferrara, 1985; Campione y Brown, 1987; Feuerstein, 1991; Resing, 1993, 2000; Jensen, 2000; Lidz y Elliot, 2000; Swanson, 2000; Guthke y Beckmann, 2000; Saldaña, 2001; Navarro, 2004).

2.2 Aproximación a las pruebas dinámicas y a los diseños metodológicos de referencia

Aunque es difícil dar una definición de las pruebas de evaluación dinámica, debido a que, como tendremos oportunidad de ver un poco más adelante, el desarrollo de éstas se ha diversificado mucho, podríamos considerarlas básicamente como actividades de evaluación que se caracterizan por la interacción entre evaluador y sujeto evaluado, y que a su vez incluyen elementos de mediación o intervención educativa en el proceso de evaluación. En general, el diseño de la mayoría de las pruebas de evaluación dinámica desarrolladas responden a dos modelos: (a) pretest/ mediación /postest y (b) pretest/ tarea  retroalimentación continua /postest. En el primero, los sujetos reciben un pretest formado habitualmente por pruebas convencionales o construidas

expresamente, pero que en cualquier caso poseen un carácter estático; posteriormente, en la fase de mediación o intervención, los sujetos reciben instrucción relacionada con las aptitudes valoradas en las pruebas de la fase pretest y, por último, se aplica un postest que normalmente está formado por las mismas pruebas que configuraron la fase inicial o pruebas equivalentes. La respuesta del sujeto a la mediación y la valoración de su desarrollo potencial vendría dado por las diferencias establecidas entre la valoración obtenida en las fases pre y postest. Según Fernández-Ballesteros (1996), una estrategia metodológica basada en un modelo pretest/ mediación /postest debería tener las siguientes propiedades: (a) evidencia de que el entrenamiento es efectivo; esto es, que tiene efectos sobre las variables dependientes; (b) evidencia de que estos efectos son perdurables; y (c) evidencia de que los cambios producidos en la situación de prueba predicen mejoras en otras situaciones de aprendizaje con respecto al funcionamiento cognitivo. Por su parte, en el segundo de los modelos más utilizados se actúa de forma que cada uno de los ítems propuestos en la prueba de evaluación va acompañado de una serie de pistas o sugerencias que, de una forma graduada, van presentándose al sujeto en el caso de que no resuelva el ítem. La cantidad y el tipo de ayuda que reciben los sujetos son determinadas por el evaluador en función de cuándo y cómo son resueltos cada uno de los ítems de respuesta. La ayuda puede continuar gradualmente hasta que se llegue a la solución correcta. Esta cantidad y tipo de ayuda recibida es lo que sirve en este modelo para determinar una valoración dinámica del desarrollo. Se intenta determinar de esta forma el tipo y la cantidad de apoyo que un sujeto necesita para resolver con éxito una determinada actividad. Aunque la mayoría de las pruebas desarrolladas se ajustan más o menos a estos dos modelos (Sternberg y Grigorenko, 2003), existen otras posibilidades de evaluación dinámica que examinaremos más adelante.

Por otra parte, existen notables e importantes diferencias entre los distintos modelos o aproximaciones de evaluación dinámica. Estas diferencias se concretan en aspectos relativos a (1) la consideración de la necesidad o no de cuantificar los resultados; (2) el tipo de información que se deriva de la aplicación de las pruebas, es decir, el planteamiento de orientaciones psicopedagógicas con base en la valoración realizada, dirigida a los educadores, o la inclusión de la predicción o el potencial de aprendizaje; (3) la población beneficiaria de la aplicación de las pruebas; (4) la naturaleza de las pruebas y de las tareas que incluyen; y (5) la naturaleza de la orientación de las pruebas, es decir, dónde focaliza su atención la actividad de evaluación (en la tarea, en el sujeto, en el contexto, en una interacción de algunos de estos factores, etc.).

Relacionado con la naturaleza de las tareas propuestas en las pruebas dinámicas, habría dos modelos u opciones posibles que tienen a nuestro entender una relevancia fundamental, sobre todo para la educación escolar. Por una parte están las pruebas que incluyen *tareas nuevas para el alumno*; estas tareas se caracterizan por su descontextualización y su semejanza con las que contienen los test convencionales, y pretenden evaluar la capacidad de aprendizaje del alumno (Feuerstein, 1991; Kozulin, 2000; Sternberg y Grigorenko, 2003). Entre sus ventajas se encontrarían: (a) evaluar el potencial de aprendizaje en tareas «libres de cultura», (b) una supuesta igualdad de condiciones al enfrentarse a las tareas y (c) evaluar aspectos generales de la inteligencia. Entre los inconvenientes podríamos destacar: (a) evaluar el potencial de aprendizaje «en abstracto», sin correspondencia con los aprendizajes tenidos en cuenta en la escuela, (b) la descontextualización de la actividad de aprendizaje, (c) la escasez o ausencia de validez ecológica y (d) un supuesto menor poder de predicción del rendimiento escolar como consecuencia de los elementos citados anteriormente.

Por otra parte, están las *tareas cuyo contenido se relaciona con los aprendizajes escolares* (Cioffi y Carney, 1983; Kletzien y Bednar, 1990; Kozulin y Garb, 2001; Grigorenko, Sternberg y Ehrman, 2000; Navarro y Mora, 2003b; Sternberg y Grigorenko, 2003). Entre sus ventajas contaríamos: (a) la contextualización de los instrumentos de evaluación, (b) un supuesto mayor poder de predicción del rendimiento escolar y (c) una mayor validez ecológica de los resultados. Entre sus inconvenientes estarían: (a) que al ser actividades conocidas, cuyo contenido está saturado de información que puede formar ya parte del bagaje de conocimientos de los sujetos, sería más difícil establecer la capacidad de aprendizaje como tal, y, como consecuencia de lo anterior, (b) la necesidad de establecer exhaustivamente el nivel inicial de desarrollo para poder determinar las mejoras.

En el siguiente cuadro podemos observar y comparar los elementos esenciales y también diferenciales de algunos modelos y pruebas de evaluación dinámica. Aunque hemos tomado como referencia la información proporcionada por Lidz y Elliot (2000), Saldaña (2001) y Sternberg y Grigorenko (2003), hemos adaptado e introducido algunos datos que creíamos relevantes para ofrecer una revisión lo más completa posible y clarificar la información.

La necesidad de una evaluación dinámica

Una de las razones fundamentales que se esgrimen para proponer procedimientos de evaluación dinámica en educación escolar es su capacidad para establecer diferencias interindividuales minimizando la influencia de la historia previa de los sujetos. En efecto, al tratarse de actividades que tratan de valorar el proceso y la capacidad de aprendizaje de determinados conceptos o habilidades, estableciendo un punto de partida común o conocido con respecto a los contenidos a trabajar, y centrandó su atención especialmente en aquellas capacidades que se encuentran en proceso de desarrollo, estos procedimientos estarían indudablemente mejor situados que las pruebas estáticas para dar cuenta de estos aspectos y muy especialmente en alumnado con privación sociocultural, dificultades de aprendizaje o necesidades educativas especiales asociadas a algún tipo de discapacidad. Este sector de la población escolar, que habitualmente muestra escaso rendimiento en pruebas de carácter convencional y estático, ha sido precisamente el principal destinatario al que se han dirigido, tanto la investigación como las aplicaciones prácticas con una metodología dinámica.

En algunos estudios se ha puesto de manifiesto que niños con muy diferentes historias de escolarización, podían igualar su rendimiento en pruebas que evaluaban su capacidad de aprendizaje (Kaniel, Tzuriel, Feuerstein, Ben-Shachar y Eitan, 1991; Feuerstein, 1991). También existen evidencias, como apuntábamos al comienzo de este capítulo, de que alumnos y alumnas con déficit cognitivo mejoraban significativamente su rendimiento en situaciones de aprendizaje asistido o en colaboración, lo que equivale a decir en situaciones de evaluación dinámica (Brown y Ferrara, 1985; Campione, Brown y Ferrara, 1987; Feuerstein, Rand y Hoffman, 1979; Feuerstein, 1991; Guthke y Beckmann, 2000; Resing, 2000; Navarro y Mora, 2003b). Por otra parte, como decíamos antes, los niños y niñas con déficit cognitivo, privación sociocultural o dificultades de aprendizaje obtienen bajas o muy bajas puntuaciones en pruebas convencionales; sin embargo no obtenemos una información relevante de éstas acerca de los procesos de aprendizaje en desarrollo, ni de cómo podríamos ayudar a estos alumnos a mejorar su rendimiento. Por tanto, la evaluación dinámica ofrecería en

Cuadro 9. Modelos y propuestas de evaluación dinámica

Enfoque y método	Población beneficiaria	Formato/diseño	Objetivo de la evaluación	Naturaleza De la tarea	Resultado y variables de medida	Estructuración de la interacción	Poder de predicción
Teoría de la EAM y la Modificabilidad estructural cognitiva. LPAD (Feuerstein y cols.)	10-16/Sujetos con desventaja sociocultural y discapacidad intelectual, ampliado a toda la población	Pretest/ mediación /postest	Producir cambios estructurales a nivel cognitivo. Establecer la modificabilidad del sujeto	Artificial Descontextualizado con respecto al programa escolar	Puntuaciones dinámicas y juicio clínico del evaluador. Se obtiene un índice de modificabilidad del sujeto	Flexible/ poco estructurada	No está bien establecido
* Batería de Modificabilidad Cognitiva (CMB) (Tzuriel)	Educación Infantil (4-5 años), alumnado de 1º ciclo de EP y alumnado mayor con NEE	Pretest/ mediación /postest	Identificación temprana de funciones cognitivas deficientes y establecer índice de modificabilidad	Artificial Descontextualizado con respecto al programa escolar	Puntuación postest y de ganancia, y valoración cualitativa	Flexible/ poco estructurada	No está bien establecido
Modelo de ED MindLadder (Jensen)	Escolares con y sin NEE	Pretest/ mediación /postest	Identificar funciones de construcción del conocimiento que se hallan en desarrollo y ofrecer pautas de intervención	Utilización de pruebas de aptitud convencionales y también de pruebas relacionadas con el currículum escolar	Análisis del cambio intrasujeto a través de la actuación en las pruebas	Estructuración alta marcada por el entorno multimedia en que se desarrolla el proceso	No está bien establecido
Modelo de ED basado en el currículum. ACFS (Lidz)	Educación Infantil (3-5 años) y alumnado mayor con NEE	Pretest/ mediación /postest	Establecer la modificabilidad y la calidad de R a la mediación	Contextualizado. Tareas relacionadas con el currículum escolar	Puntuaciones dinámicas y valoración cualitativa	Flexible/ semi-estructurada	No está establecido
Enfoque de sugerencias graduadas (Campione y Brown)	Alumnado con bajo rendimiento académico	pretest/ tarea retroalimentación continua /postest	Determinar cantidad y tipo de apoyo necesario para lograr mejoras en aprendizaje y transferencia	Utilización de pruebas de aptitud convencionales y también de pruebas relacionadas con el currículum escolar	Cantidad y tipo de sugerencias necesarias. Se obtiene una medida del potencial de aprendizaje	Muy estructurada	No está bien establecido
Evaluación del potencial de aprendizaje por razonamiento inductivo (LIR) (Resing)	Alumnado de 7 y 8 años	Pretest/ mediación /postest	Establecer el potencial de aprendizaje mediante tareas de razonamiento inductivo	Artificial Descontextualizado con respecto al programa escolar	Cantidad y tipo de sugerencias necesarias. Resultado contemplado como información adicional a la de los test convencionales	Muy estructurada	No está bien establecido

Cuadro 9. (Continuación)

Prueba del Potencial de Aprendizaje (Budoff)	Sujetos con bajo rendimiento en pruebas de CI que han experimentado fracaso escolar	Pretest/ mediación /postest	Ofrecer un índice válido del potencial de aprendizaje y contribuir a una categorización acorde a este potencial	Artificial Descontextualizado con respecto al programa escolar	Puntuaciones postinstrucción y puntuación residual	Estructuración y normalización aproximadas	Alto con respecto a la población designada
Evaluación del Potencial de Aprendizaje (EPA) (Fernández-Ballesteros y cols.)	Niños y niñas a partir de 5 años, alumnado con NEE y mayores	Pretest/ mediación /postest	Establecer un índice válido del potencial de aprendizaje	Artificial Descontextualizado con respecto al programa escolar	Puntuación de mejora como medida del potencial de aprendizaje, y valoración cualitativa	Muy estructurada	Alto con respecto al constructo entrenado en la fase de intervención
Enfoque Lerntest (Pruebas de Aprendizaje) (Guthke y cols.)	Alumnado con NEE y adultos con lesión cerebral	Pretest/ mediación /postest (prolongada) e instrucción dentro del procedimiento de prueba (breve)	Evaluación del proceso y la capacidad de aprendizaje	Utilización de pruebas descontextualizadas y también de otras relacionadas con el currículum escolar	Puntuación postest como medida del potencial de aprendizaje y valoración cualitativa	Muy estructurada y normalizada	Muy alto para sujetos con bajas puntuaciones en pruebas de CI
Enfoque del Procesamiento de la información (TPC-S) (Swanson)	Alumnado con DA y NEE	Pretest/ mediación /postest	Evaluar dinámicamente la MT y ofrecer un indicador del potencial de procesamiento	Tareas específicas de MT Descontextualizado con respecto al programa escolar	Siete puntuaciones sirven para calcular la medida del potencial de procesamiento (mejora, Mantenimiento, diferencia de procesamiento...)	Muy estructurada y normalizada	Alto para sujetos con DA
Enfoque de Prueba de los Límites (Carlson y Wiedl)	Alumnado con y sin DA y NEE	Distintas condiciones de intervención (cantidad variable de verbalización y retroalimentación)	Establecer las condiciones óptimas para la mejora del rendimiento en la prueba	Artificial Descontextualizado del programa escolar	Puntuaciones obtenidas en las distintas condiciones de prueba evaluadas. Determinación de la condición óptima	Muy estructurada y normalizada	Predicción del rendimiento escolar similar a la de pruebas de CI y mayor ante métodos dinámicos de enseñanza
Evaluación Dinámica de Procesos Metacognitivos (Saldaña)	Sujetos con bajo nivel de funcionamiento cognitivo	Interacción- Retroalimentación continua 	Evaluar dinámicamente procesos metacognitivos (metaconocimientos y procesos de autorregulación)	Utilización de pruebas convencionales y descontextualizadas, y también de actividades relacionadas con el contexto de aplicación	Cantidad y tipo de ayuda necesaria	Muy estructurada	Limitada por los estudios realizados. Predicción relativa de las conductas de adaptación social y de habilidades metacognitivas en procesos de intervención

principio una situación de valoración más justa o equitativa para los sujetos que afrontan una prueba con desventaja, ya sea de tipo social, cultural o cognitiva; fundamentalmente, al evaluar el desarrollo potencial de su aprendizaje, igualando, en cierto modo, las condiciones de partida.

Así mismo, mediante la evaluación dinámica podemos acceder a la valoración de distintos procesos o elementos implicados en una actividad determinada; la información sobre los mecanismos a través de los cuales un alumno aprende un determinado contenido, puede ser de suma importancia a la hora de orientar la intervención. De esta forma, se establece una conexión real entre evaluación e intervención, lo que ha sido destacado por diversas voces como clave para mejorar los procesos de enseñanza y aprendizaje (Vygotski, 1934/1995, 1979/1996; Solè, 1992; Maruny, Ministrál y Mirallés, 2001; Coll, Martín y Onrubia, 2001; Coll, Barberá y Onrubia, 2000; Sternberg y Grigorenko, 2003; Aguilera y García, 2004).

2.3 Modelos y propuestas de evaluación dinámica

Introducción

En páginas anteriores hemos visto cómo se desarrolló originalmente el concepto de evaluación dinámica, así como algunos de los principios teóricos y metodológicos que han sido claves en este desarrollo. Desde que Vygotski, primero, y Feuerstein, posteriormente, realizaran sus propuestas de evaluación partiendo de una concepción dinámica del desarrollo y el aprendizaje, han tenido lugar no pocos estudios que han tratado básicamente de mostrar la idoneidad de una aproximación dinámica o interactiva en la valoración de las funciones cognitivas. Por una parte, y más específicamente, estos estudios han tratado de desarrollar las ideas e intuiciones del psicólogo ruso acerca de la evaluación de y en la zona de desarrollo potencial; por otra, un número importante de investigaciones, enraizadas en los trabajos pioneros de Feuerstein sobre la evaluación del potencial de aprendizaje, se han dirigido a replicar, desarrollar y completar los estudios y la teoría de Feuerstein sobre la EAM y la modificabilidad estructural cognitiva (Alonso Tapia, 1992; Kozulin, 2000). El contenido de las páginas que siguen está dedicado precisamente a exponer una revisión de algunos de estos modelos y enfoques dinámicos de evaluación. Sin embargo, tal y como apuntábamos anteriormente, a lo largo de las últimas décadas se han ido desarrollando y diversificando nuevos enfoques de evaluación dinámica. Algunos constituyen en realidad aplicaciones adaptadas de los primeros enfoques dinámicos a diferentes realidades y situaciones; otros en cambio, se sustentan sobre nuevas conceptualizaciones de la ZDP o sobre la constatación empírica de la relevancia de determinados procesos cognitivos para el aprendizaje y han ido construyendo una sólida base teórica. La mayoría sigue apostando por la valoración de funciones o procesos generales de aprendizaje, pero cada vez más las aplicaciones dinámicas parecen interesarse por procesos cognitivos concretos o incluso por dominios específicos de áreas de aprendizaje. Lógicamente no están todos, eso no sería posible, pero hemos procurado analizar, —con mayor o menor profundidad en función de la información lograda— las características esenciales de los modelos que han contribuido de una forma especial al estudio y aplicación de la evaluación dinámica. También hemos considerado aquellos enfoques que, aún no teniendo de momento una gran repercusión en el ámbito científico internacional, o bien, aún tratándose de adaptaciones de modelos

dinámicos clásicos, respondían a un interés particular relacionado con algunas de las características de nuestro propio estudio.

Por otra parte, hemos prestado una especial atención a aquellos enfoques o estudios que, de una u otra forma, han considerado el lenguaje escrito entre los objetivos y contenidos de sus propuestas. Anticipamos que no ha sido fácil encontrarlos, ya que, como recogíamos con anterioridad, la gran mayoría de las aplicaciones dinámicas se ha efectuado con tareas provenientes de pruebas convencionales de la inteligencia, o adaptaciones de las mismas, considerando, como tendremos oportunidad de comprobar, que solo el uso de tareas «libres de influencias culturales» podría ofrecer la posibilidad de valorar el potencial de aprendizaje en igualdad de condiciones para sujetos con historias distintas de aprendizaje. Por otra parte, algunos estudios tratarán de revelar el poder de predicción de las pruebas dinámicas sobre el rendimiento escolar en lectura o sobre situaciones de aprendizaje prolongadas; es en estos casos, cuando podremos al menos intuir las consecuencias que para el logro académico en lectura y comprensión pueden tener determinadas formas de mediación aplicadas en el proceso de evaluación dinámica, así como el contenido específico de la mediación realizada.

En el análisis de cada uno de los modelos o enfoques presentados se distinguen dos apartados. En el primero de ellos, *Bases teóricas y metodológicas*, se ha tratado de destacar la teoría de base que sustenta la propuesta, así como el diseño metodológico propuesto para valorar el potencial de aprendizaje. También se ha procurado obtener información sobre la población beneficiaria y la naturaleza de las tareas que conforman las pruebas dinámicas utilizadas. En el segundo de los apartados, *Investigación y revisión crítica del modelo*, se recogen algunos datos sobre los estudios más significativos que se han realizado, así como sobre sus posibles implicaciones para la evaluación educativa. Igualmente, de forma transversal a todos estos elementos que caracterizan a las distintas aproximaciones dinámicas, pero especialmente unidos a los estudios presentados, planteamos algunos análisis críticos con relación a los mismos y al modelo propuesto; en ocasiones se tratará de objeciones, bien recogidas por otros autores o bien realizadas por nosotros mismos, y otras veces se tratará de comentarios críticos que respaldan o amplían los presupuestos o las conclusiones de los enfoques y de los estudios. En ambos casos, objeciones o acuerdo, trataremos de enlazar estos análisis con nuestra posterior propuesta de evaluación dinámica de procesos implicados en la lectura, así como con nuestra discusión del proceso y de los resultados obtenidos.

La evaluación del Potencial de Aprendizaje de R. Feuerstein

Bases teóricas y metodológicas

La evaluación dinámica propuesta por Feuerstein, concretada en el dispositivo de evaluación del potencial de aprendizaje (LPAD son sus siglas en inglés), es un método de evaluación del potencial de crecimiento cognitivo en niños, adolescentes y adultos. Lo que se trata de poner de manifiesto durante el proceso de valoración es el grado de modificabilidad que posee un sujeto (Feuerstein, 1991; Feuerstein et. al., 1995). Esta estrategia de evaluación se basa en un modelo pretest/mediación/posttest y consta de dos momentos diferenciados. En un primer momento se evalúan una serie de habilidades de pensamiento, procesos psicológicos básicos, aprendizaje y resolución de problemas; posteriormente, durante la fase de mediación tiene lugar una intervención cuidadosamente organizada, pero estructurada flexiblemente, en la que se abordan las

habilidades y los principios cognitivos necesarios para resolver con éxito las tareas propuestas anteriormente, seguida o acompañada de una valoración de la medida en que la actividad desarrollada es capaz de modificar al sujeto en la dirección de ver incrementada su capacidad para resolver los problemas planteados. Los objetivos de este procedimiento de evaluación incluyen: (a) identificar funciones cognitivas bien desarrolladas, (b) identificar funciones cognitivas deficientes, (c) estimar la modalidad y la cantidad de mediación necesaria para superar las deficiencias cognitivas manifestadas y (d) dirigir la atención del evaluador y del sujeto hacia los procesos implicados en la resolución de las tareas planteadas.

Ya realizamos con anterioridad una exposición de algunas premisas teóricas y metodológicas en las que se basaba el trabajo desarrollado por Feuerstein y sus colaboradores. Revisando las ideas fundamentales del concepto de mediación, nos deteníamos en la teoría de la *experiencia de aprendizaje mediado* (EAM) y la modificabilidad estructural cognitiva, en la cual hunde sus raíces el instrumento que ahora analizamos. Según Feuerstein, la EAM, como proceso universal y factor decisivo y explicativo del desarrollo humano, tiene varios componentes que una vez «destilados» pueden convertirse en la esencia de una acción educativa intencional para mejorar un desarrollo cognitivo deficiente y, así mismo, un proceso educativo de evaluación. Estos componentes serían: (a) la selección de estímulos, (b) la organización de estos estímulos, (c) la previsión o anticipación de resultados, (d) la presentación de modelos a imitar, (e) la presentación de estímulos culturalmente específicos, (f) la repetición y la diversificación, (g) la transmisión cultural del pasado y la representación del futuro, y (h) la conducta comparativa.

Entre las aportaciones de la teoría de la EAM de Reuven Feuerstein y sus colaboradores, una de las más relevantes y que, de una u otra forma ha sido asumida por los investigadores de la evaluación dinámica, ha sido la de presuponer en colectivos no alfabetizados, pero que contaban con una tradición cultural de transmisión oral a través de la cual tenían lugar las experiencias mediadas de aprendizaje, un alto potencial de aprendizaje, que terminaba por equiparar estos grupos, tras un proceso de evaluación e intervención adecuados, al resto de la población de condiciones socioculturales normalizadas (Feuerstein, 1991, 1998a, 1998b; Kaniel, Tzuriel, Feuerstein, Ben-Shachar y Eitan, 1991). Esta hipótesis ha dado origen a numerosos e interesantes estudios realizados con población culturalmente diversa que, por diversas circunstancias históricas y sociales se encontraban en una situación de desventaja social (Emerson, 1985, 1986; Kozulin, 2000).

Feuerstein propuso también varias funciones cognitivas deficientes (cuadro 10), que podían ser el resultado de la ausencia de EAM (cuadro 8) y que agrupó en cuatro categorías: (1) problemas cognitivos en la fase de *input*, (2) problemas cognitivos en la fase de elaboración, (3) problemas de cognición en la fase de *output*, y (4) factores motivadores de carácter afectivo. (En realidad esta última categoría fue incluida con posterioridad y carece de unas funciones concretas asignadas; parece tratarse en realidad de un reconocimiento de su relevancia, pero sin que tengamos constancia de que se haya tratado de operativizar en modo alguno).

Cuadro 10. Funciones cognitivas deficientes. Tomado de: El modelo de Experiencia de Aprendizaje Mediado (EAM). Jerusalén, 1998

Fase de input:

1. Percepción borrosa y confusa
2. Comportamiento exploratorio impulsivo, asistemático, no planificado
3. Falta de instrumentos verbales y conceptos para identificar los objetos
4. Orientación espacial deficiente
5. Orientación temporal deficiente
6. Deficiencia en la constancia y permanencia del objeto
7. Deficiencia en la precisión y exactitud en la recopilación de datos
8. Dificultad para considerar dos o más fuentes de información a la vez

Fase de elaboración:

1. Dificultad para percibir un problema y definirlo
2. Dificultad para distinguir los datos relevantes de los irrelevantes
3. Dificultad o carencia de conducta comparativa
4. Estrechez del campo mental
5. Percepción episódica de la realidad
6. Carencia de razonamiento lógico
7. Carencia de interiorización del propio comportamiento
8. Restricción del pensamiento hipotético inferencial
9. Carencia de estrategias para verificar hipótesis
10. Dificultad en la planificación de la conducta
11. Dificultad en la elaboración de categorías cognitivas
12. Dificultad para la conducta sumativa
13. Dificultad para establecer relaciones virtuales

Fase de output:

1. Comunicación egocéntrica
2. Dificultad para proyectar relaciones virtuales
3. Bloqueo en la comunicación de la respuesta
4. Respuestas por ensayo-error
5. Carencia de instrumentos verbales adecuados
6. Carencia de la necesidad de precisión y exactitud para comunicar las propias contestaciones
7. Deficiencia en el transporte visual
8. Conducta impulsiva

El LPAD se centra en una evaluación del proceso de cambio, no en la valoración de un producto (Feuerstein, 1991). Este dispositivo se utiliza para responder a una serie de preguntas relacionadas con el sujeto evaluado, las cuales deben orientar al evaluador durante el proceso de evaluación: ¿es el individuo tan modificable como afirma el postulado general?, ¿son los distintos niveles de la modificabilidad dependientes de la condición individual, de la línea base establecida y de una variedad de otros determinantes, tales como la intensidad de la EAM a la que ha sido expuesto el individuo?, ¿cuál es la naturaleza y el grado de cambio que uno puede esperar?, ¿cuál es la naturaleza y la cantidad de mediación necesaria para conseguir que los cambios deseados sean permanentes y a largo plazo?. Por otra parte, el LPAD está orientado a

establecer el potencial de aprendizaje y la modificabilidad de un sujeto, no en un sentido estático, en el que se asume un determinado límite al determinar un potencial de desarrollo, sino en un sentido dinámico, asumiendo la expansión de los límites en función de la interiorización de las nuevas capacidades (Feuerstein et al., 1995). También pretende determinar u orientar la modalidad —verbal, visual, motriz, etc.— preferente a través de la cual la modificabilidad del sujeto puede tener lugar de una manera óptima.

El dispositivo de evaluación del potencial de aprendizaje de Feuerstein consta de 15 instrumentos, la mayoría de los cuales están basados en pruebas psicométricas normalizadas (Grigorenko y Sternberg, 1998; Sternberg y Grigorenko, 2003; Alonso Tapia, 1992). La metodología aplicada al uso de estos instrumentos es lo que introduce la diferencia fundamental con la aplicación estándar de las pruebas originales. Algunos instrumentos del LPAD pueden ser aplicados grupalmente, mientras que otros requieren una administración individual; en el cuadro 11 hemos recogido el total de los instrumentos que componen el dispositivo, distinguiéndolos en función de esta característica. En general, la administración del LPAD no se suele hacer de forma completa, lo cual sería muy costoso en tiempo y energías, tanto para el evaluador como para el sujeto evaluado (Tzuriel, 1995a). Suelen emplearse unas pruebas u otras en función del perfil individual, y la cantidad de mediación necesaria suele ir determinando la duración de la aplicación (Kozulin y Falik, 1995; Feuerstein et al., 1995). El dispositivo tuvo su origen, y ha sido tradicionalmente empleado, en el trabajo con población infantil y juvenil (de los 10 a los 16 años) con discapacidad, que obtenían bajas puntuaciones en test convencionales de inteligencia (Feuerstein, Rand y Rynders, 1988; Feuerstein, 1998a) y también con niños y niñas pertenecientes a minorías étnicas o deprivados socioculturalmente (Kaniel, Tzuriel, Feuerstein, Ben-Shachar y Eitan, 1991; Feuerstein, 1991), aunque en los últimos años se está redefiniendo la población diana de este dispositivo para incluir a niños sin dificultades en su desarrollo ni en su aprendizaje (Kozulin, 2000; Sternberg y Grigorenko, 2003).

Cuadro 11. Instrumentos que conforman el dispositivo de evaluación del potencial de aprendizaje (LPAD)

Aplicación individual y en grupo	Aplicación individual
<ul style="list-style-type: none"> *Organización de puntos *Matrices progresivas coloreadas de Raven *LPAD Set Variaciones B8/B12 *LPAD Set Variaciones I *LPAD Set Variaciones II *RSDT (Representational Stencil Design Test: Prueba de diseño de patrones) *Progresiones numéricas *Figura compleja de Rey-Osterre *PLT (Positional Learning Test: Prueba de aprendizaje de posiciones 5x25) *Organizador 	<ul style="list-style-type: none"> *Lahi Test (Mantenimiento de la atención) *Test de recuerdo asociativo de estímulos visuales I: Reducción funcional *Test de recuerdo asociativo de estímulos visuales II: Parte-todo *Plataforma *Test de recuerdo de palabras

La aplicación del LPAD ofrecería la posibilidad de establecer una experiencia de aprendizaje mediado mediante la creación de una zona de desarrollo próximo (Sternberg y Grigorenko, 2003). Esta interpretación del concepto es, sin duda, una de las más relevantes desde nuestro punto de vista y es aquí, donde las confluencias teóricas entre las ideas originarias de Vygotski y la formulación y concreción de la teoría de la EAM por parte de Feuerstein, se hacen especialmente palpables. De igual relevancia es, en este sentido, que el componente afectivo de la interacción se torne especialmente importante; en efecto, en contraste con las situaciones de evaluación estática, en las que se destaca la importancia de que el evaluador mantenga una actitud neutral e incluso «aséptica» durante la administración de las pruebas, en la evaluación dinámica propuesta por Feuerstein, la relación afectiva entre el evaluador y el sujeto es parte integrante de la situación de prueba y se manifiesta en la atención y supervisión continua del proceso y en la respuesta del evaluador a las preguntas y comentarios planteados durante la aplicación, mostrando abiertamente interés y manteniendo una actitud de diálogo e interacción continua.

El modelo de Feuerstein propone que los cambios que pueden observarse durante la aplicación del LPAD orientan al evaluador acerca del tipo y la cantidad de intervención necesaria para producir transformaciones significativas a nivel de estructuras cognitivas. La calidad de la respuesta del niño a la mediación se operativiza mediante el concepto al que ya hemos aludido de modificabilidad estructural cognitiva, pudiéndose obtener una puntuación con base en la administración de las pruebas, así como en el juicio clínico del evaluador. Habitualmente el resultado de la evaluación conlleva unas orientaciones sobre la intervención óptima que debe recibir el niño para ver incrementado su potencial de aprendizaje.

Investigación y revisión crítica del modelo

De entre las múltiples investigaciones que se han desarrollado en el marco de las propuestas de evaluación del potencial de aprendizaje de Feuerstein y mediante la utilización de instrumentos pertenecientes al LPAD, hemos seleccionado una pequeña muestra que podría representar al menos los centros de interés que han guiado principalmente estos estudios, así como algunos aspectos que nos resultan particularmente relevantes con relación a nuestro propio trabajo. Por una parte, aunque no ha sido objeto de atención en la mayoría de las investigaciones publicadas, un tema que consideramos central en la propuesta de Feuerstein es la fiabilidad de la fase de mediación en el proceso de evaluación, ya que, como decíamos en páginas anteriores, se trata de una intervención en la que el evaluador juega un papel fundamental, con un marcado carácter clínico (Lidz, 1991), que dificulta, por una parte, la replicabilidad de las observaciones y de las valoraciones efectuadas y exige, por tanto, una formación exhaustiva a los aplicadores del dispositivo (Sternberg y Grigorenko, 2003). En palabras de estos autores, los bajos índices de fiabilidad entre observadores presentados, podrían llegar a ser más indicativos de las características del evaluador que de las del niño. En este sentido, algunos estudios han revelado que la fiabilidad entre distintos observadores no ha alcanzado niveles del todo aceptables, incluso cuando se trataba de valorar la presencia de determinadas funciones cognitivas deficientes del mapa cognitivo elaborado desde la propia teoría. Esta fiabilidad oscilaba entre niveles del 31% y el 97,1% con una mediana del 65,7% en un estudio realizado para evaluar el acuerdo entre observadores con una escala de observación de las funciones mencionadas (Tzuriel y Samuels, 2000).

Una serie de estudios muy interesantes ha sido el que se ha dedicado a comprobar la validez de diversos tipos de mediación en la ejecución de las diversas pruebas que componen el dispositivo. En un estudio de Tzuriel y Feuerstein (1992) se comparó la actuación en el test de matrices progresivas estándar de Raven (Raven, 1995) con una versión adaptada de este mismo test y aplicada de forma dinámica que forma parte de los instrumentos del LPAD (Set Variaciones II). El total de la amplia muestra era de 1394 niños de cuarto a noveno curso, pertenecientes a entornos desfavorecidos y privilegiados. La aplicación del subtest citado se realizó en grupo y constó de una fase de demostración y posteriormente las tres fases clásicas en la aplicación de la evaluación dinámica: pretest/mediación/postest. El investigador introdujo la variable «cantidad de mediación aportada», aportando mucha, poca o ninguna intervención. Por otra parte, la administración del test de matrices progresivas se realizó dos veces, antes de la intervención y dos semanas después. El diseño contempló un grupo control con relación a la administración del test de matrices, es decir, un grupo que no había recibido ninguna mediación; sin embargo, no se contempló la existencia de grupo control con respecto a la aplicación de la prueba dinámica. Los resultados revelaron varios datos de interés: en primer lugar, los efectos de la prueba dinámica se podían predecir por los resultados obtenidos en el test de matrices aplicado de forma convencional; en segundo lugar, los chicos de entornos privilegiados obtuvieron mejores resultados que los desfavorecidos en ambas pruebas, sin embargo, la curva de resultados tendía a igualarse con respecto al grupo de niños desfavorecidos que había recibido una mayor cantidad de mediación; en tercer lugar, el efecto de la cantidad de mediación fue significativo en general; la cuarta consideración que realizan los autores hace referencia a que los resultados en el postest de la prueba dinámica fueron significativamente mejores que los del pretest; y, por último, hacen referencia a que los efectos observados en el postest de la prueba dinámica persistían dos semanas más tarde cuando fue aplicado de nuevo el test de Raven.

A pesar del interés de éste y de otros estudios similares con relación a determinar la relevancia del tipo y cantidad de mediación realizada en la mejora del rendimiento en las pruebas, por una parte; los mejores resultados obtenidos por estudiantes en situación de desventaja sociocultural cuando son evaluados dinámicamente, por otra; así como a establecer diferencias entre actuaciones previas y posteriores a la mediación, varios son los análisis desde una perspectiva crítica que han acentuado especialmente los problemas metodológicos de la mayoría de estas investigaciones. Haremos mención particularmente a dos aspectos recogidos en la revisión de Sternberg y Grigorenko (2003). En primer lugar, la no existencia de grupo control (grupo que no recibía mediación) en la aplicación de la prueba dinámica, introduce un sesgo importante por cuanto se produce un efecto de interacción de las aptitudes evaluadas con el tratamiento proporcionado. La utilización e incorporación al modelo de una medida del potencial de aprendizaje al grupo que no recibía mediación alguna, hubiera permitido controlar mejor el efecto de la cantidad de mediación aportada en la aplicación del subtest del LPAD utilizado; de otra forma, los resultados pueden ser limitados al revelar que mayores cantidades de intervención conducen a mejores resultados y que niños con un nivel inicial bajo mostraron mejoras cuando eran instruidos específicamente en la prueba sobre las que se les iba a evaluar.

Relacionado en parte con lo anteriormente expuesto, la segunda cuestión relevante en el análisis de estas investigaciones, se refiere al problema del posible efecto de la práctica. Muchos son los autores y los estudios que han advertido sobre lo dudoso de

determinadas conclusiones que no tienen en cuenta este efecto (Mora, 1986; Serpell, 1993). De este modo, se aduce que los resultados aportados por la investigación relatada más arriba, podrían ser obtenidos en diseños similares utilizando pruebas estáticas, ya que la mera repetición de la prueba origina incrementos que pueden llegar a ser significativos en algunos casos.

Como decíamos al comienzo de este análisis sobre los estudios desarrollados sobre este modelo, muchos se han desarrollado con población inmigrante en situación de desventaja social. En un trabajo llevado a cabo con población etíope que emigró a Israel durante los años ochenta (Kaniel, Tzuriel, Feuerstein, Ben-Shachar y Eitan, 1991), se pretendía demostrar que los grupos socioculturalmente deprivados y, por consiguiente, con inadecuadas o escasas experiencias de aprendizaje mediado, obtienen un bajo rendimiento escolar y unos índices bajos de modificabilidad cognitiva en su incorporación a un nuevo entorno; y que, sin embargo, los grupos culturalmente diversos, pero que por circunstancias histórico-culturales han mantenido estructuras de transmisión cultural que permitían la existencia de adecuadas experiencias de aprendizaje mediado, obtendrían mejores niveles de adaptación escolar y mayores índices de modificabilidad. Se aplicaron varios subtest del LPAD (Matrices Progresivas de Raven: set de variaciones I y II; Organización de puntos; Figuras complejas; Organizador; Progresiones numéricas y una prueba de Progresiones figurales). El total de la muestra estaba compuesto por 373 adolescentes, 316 formaron parte del grupo experimental y 57 del grupo control. Cada una de las tareas fue aplicada en las tres fases de pretest, mediación y postest, excepto el test de Raven, que fue aplicado solo en las fases inicial y final. El GC recibió las mismas pruebas y en las mismas condiciones, pero no fue instruido en la fase de mediación. Los resultados pusieron de manifiesto que el GE se benefició de la mediación recibida en todas las tareas trabajadas, obteniendo puntuaciones en el postest que diferían significativamente de las obtenidas por el GC. Al mismo tiempo, se observó que la aplicación dinámica había introducido modificaciones en la curva de resultados esperable a tenor de los resultados del pretest; de esta forma, el pretest no podía predecir la actuación de los sujetos del GE después de la mediación. Por otra parte, estos resultados fueron comparados a su vez con los obtenidos por grupos de chicos y chicas israelíes en situación de deprivación sociocultural. En este caso, los resultados de la comparación tras la aplicación dinámica de las pruebas de Variaciones I y II, revelaron que el grupo «culturalmente diferente», es decir, el grupo de adolescentes inmigrantes etíopes, obtuvo mejores indicadores de modificabilidad que el grupo de alumnos deprivados socioculturalmente. El incremento de las puntuaciones se produce en ambos grupos, pero mientras el grupo de inmigrantes lo hace en 14 puntos de variación entre el pretest y postest del Raven (de 22 a 36), el grupo de deprivación incrementó su puntuación en 5 (de 39 a 44). También se compararon los resultados obtenidos por el GE tras la mediación, con los de la población normalizada de adolescentes israelíes, concluyendo que las diferencias previas entre ambos grupos de población tienden a disminuir de forma significativa. Finalmente, los autores realizan un análisis cualitativo de especial interés en función de los perfiles surgidos tras la evaluación dinámica. Así, distinguen entre sujetos con alto y bajo potencial de aprendizaje, basándose en esta distinción para efectuar determinadas orientaciones para una posterior intervención educativa continuada y específica.

Tras el análisis de este estudio, a todas luces revelador del interés primordial, manifestado en numerosas ocasiones por Feuerstein y compartido por la gran mayoría de los investigadores de evaluación dinámica, en mostrar la influencia decisiva de la

oportunidad de haber contado con adecuadas experiencias de aprendizaje mediado y, por tanto, con un conducto válido de transmisión cultural, en la capacidad de aprendizaje y modificabilidad de los individuos; es necesario, sin embargo, apuntar algunas objeciones que nos parecen importantes. Por una parte, una objeción, que tendremos oportunidad de volver a analizar en otros estudios, se refiere al hecho de que los autores concluyan que el GE obtuvo mejoras en su capacidad de aprendizaje — mediante el indicador de su modificabilidad— y también de transferencia, al obtener estas mejoras mediante la realización del test de Raven, en el cual no habían sido «directamente» entrenados. Esta conclusión nos parece especialmente sobrevalorada por dos razones: la primera se debe a que a los sujetos les había sido administrado el test de Raven ya en el pretest —aunque en el estudio no se hace alusión al periodo transcurrido entre las fases pre y postest, es de suponer en función de las tareas aplicadas en la fase de mediación, que se trató de un corto espacio de tiempo—, lo que podría explicar una parte de los resultados obtenidos y, lo más importante en este caso, reduciría ostensiblemente la posibilidad de interpretar que se produce una transferencia del aprendizaje efectuado; por otra parte, la segunda razón se debe precisamente al entrenamiento realizado. Si se conoce el contenido de las actividades de las pruebas Set Variaciones I y II, no es fácil admitir que no se está entrenando más o menos directamente, no solo en los principios cognitivos que presiden la ejecución del Raven, sino que se hace con tareas muy similares. Por tanto, bajo nuestro punto de vista, podríamos hablar de una transferencia de los aprendizajes muy limitada y cercana, efectuada a tareas muy similares a aquellas donde tuvieron lugar dichos aprendizajes.

Una segunda objeción tiene que ver con la comparación efectuada finalmente con los alumnos deprivados y con la población normalizada israelí utilizando las puntuaciones en el test de Raven. En esta comparación se observa un incremento muy significativo del grupo inmigrantes, siendo este sin duda el dato relevante; pero al observar los resultados del grupo deprivación notamos que estos parten con una puntuación inicial que supera significativamente a la del otro grupo y que incrementa de 39 puntos hasta situarse en 44; pues bien, esta puntuación postest coincide con la puntuación media de la población normalizada de adolescentes israelíes en el pretest. En consecuencia, no podríamos decir que los indicadores de modificabilidad son superiores en un grupo que en otro, ya que, aunque con menor diferencia, el grupo deprivación incrementa su puntuación tras la aplicación dinámica hasta situarse a un nivel de ejecución medio (percentil 50). A nuestro juicio, estos datos son muy importantes, ya que revelan ganancias significativas tanto para el grupo inmigrantes como para el grupo deprivación, pero la interpretación que realizan los autores del estudio, basada en la mayor modificabilidad del grupo inmigrantes que, a su vez, se conectaba con la tesis principal del estudio, no queda del todo clara con estas comparaciones.

Por otra parte, se han realizado otras observaciones al modelo de evaluación dinámica propuesto por Feuerstein. En primer lugar, como ya comentamos, la ampliación de la población diana hasta incluir a los sujetos que no presentan ningún tipo de dificultad, parece no estar bien justificada (Sternberg y Grigorenko, 2003). La idea fundamental de que la EAM es el factor clave que explica el desarrollo humano y la transmisión de la cultura no parece conllevar el hecho de que un instrumento de evaluación construido con base en esa teoría posea un mismo poder explicativo para distintos segmentos de población que difieren significativamente en elementos valorados en la misma prueba. En cualquier caso, según los autores citados anteriormente, esta ampliación del campo de aplicación no queda clara a la luz de los datos ofrecidos hasta ahora por las

investigaciones. En segunda instancia, los datos empíricos ofrecidos por los estudios realizados para validar el LPAD presentan, según muchos autores, poca solidez (Mora, 1986, 1998; Bradley, 1983; Grigorenko y Sternberg, 1998; Sternberg y Grigorenko, 2003; Büchel y Scharnhorst, 1993; Saldaña, 2001). Por ejemplo, Büchel y Scharnhorst han hecho hincapié en la inadecuada e imprecisa definición de los conceptos que se manejan en la teoría de la EAM y la modificabilidad estructural cognitiva y que, como consecuencia, afectan notablemente a la posibilidad de operativizar, ya sea cuantitativa o cualitativamente los resultados obtenidos en la evaluación. Esta definición vaga de conceptos relacionados con el mapa cognitivo de funciones deficientes o con las características de la EAM, hacen también difícil establecer la validez de los resultados. Feuerstein y sus colaboradores no han mostrado un interés primordial por establecer validaciones de constructo o de criterio. Esto se explica en parte debido a que, desde su punto de vista, el rendimiento escolar no sería el criterio externo más adecuado para valorar la validez predictiva del LPAD (Feuerstein, Rand y Hoffman, 1979), centrando su atención en la modificación observada en la conducta del sujeto tras la evaluación dinámica. Para validar la evaluación dinámica debería tener lugar una intervención específica orientada en la dirección apuntada por la propia evaluación (Tzuriel, 1992).

Nos gustaría hacer unas últimas referencias críticas a la propuesta desarrollada por Feuerstein, y decimos «desarrollada» y no tanto «formulada», ya que, si bien es cierto que revisando sus postulados teóricos relativos a las características de la EAM o a la configuración del mapa cognitivo de las funciones deficientes, podemos hallar evidencias claras de la relevancia de los procesos metacognitivos, a pesar de que no esté explícitamente detallado; y que no es menos cierto que últimamente y a menudo puede leerse en revisiones sobre su obra o en los artículos derivados de los estudios presentados por su grupo de colaboradores, que tanto la teoría de la EAM como los procesos de evaluación dinámica (LPAD) e intervención que surgieron en el marco de ésta, actúan sobre estos procesos de control y regulación consciente del pensamiento (Kozulin, 2000; Sternberg y Grigorenko, 2003), lo cierto es que nuestra propia experiencia directa en talleres de trabajo junto a Feuerstein y sus colaboradores y el conocimiento de la aplicación práctica de al menos algunas de las pruebas dinámicas utilizadas en el LPAD, no nos han revelado su utilización en el proceso de evaluación, ni su especial relevancia en la explicación de los resultados obtenidos (Navarro, 2004). Por otra parte, quizás sea más objetivo revisar los estudios presentados aquí mismo desde esta perspectiva, para comprobar que la alusión a componentes metacognitivos son muy tangenciales y en absoluto guían el proceso de búsqueda de posibles problemas de aprendizaje, ni están en la base de las explicaciones que se derivan de estos estudios. Aunque nos detendremos posteriormente de forma especial en la relevancia de los procesos metacognitivos para poder obtener información valiosa sobre el proceso de aprendizaje, cuando abordemos nuestra propuesta, por ahora sólo destacar la posible relación que puede existir entre la falta de validez predictiva que parecen haber apuntado los estudios con el LPAD y la ausencia —al menos explícita— de un análisis de los procesos metacognitivos que sin duda se ponen de manifiesto en el transcurso de las tareas planteadas.

Por otra parte, ya hemos hecho alusión de algún modo a otro de los aspectos que nos interesan especialmente, este es la contextualización o descontextualización de los instrumentos de evaluación. El modelo de evaluación dinámica de Feuerstein, o mejor dicho, su concreción en el LPAD, es quizás el paradigma de prueba dinámica descontextualizada. Aunque es necesario recordar que el proceso de interacción

mediada, característico de la evaluación dinámica, es un elemento que de por sí aporta un grado de contextualización en el proceso de evaluación, ya que habitualmente todos los procesos de aprendizaje tienen lugar en contextos de interacción; lo cierto es que, una vez valorado este hecho, tanto si nos referimos al contexto escolar, como si hacemos referencia a un contexto social, la mayoría de los instrumentos que contiene el LPAD, así como su aplicación práctica, no guarda relación significativa con los contenidos ni con el contexto de referencia, en el cual supuestamente deben ser puestas en práctica las habilidades aprendidas durante el proceso. Como apuntábamos anteriormente, no hay prácticamente ningún estudio que aplique un criterio externo que de alguna forma «valide» las mediciones realizadas sobre las funciones cognitivas modificadas (Sternberg y Grigorenko, 2003).

Por último, destacaremos un aspecto que ha ido cobrando progresivamente interés dentro de la teoría de Feuerstein. Nos referimos a la inclusión del contexto en el esquema interpretativo y explicativo del desarrollo y, por tanto, de la intervención y la evaluación. Hasta el momento, la constatación de la relevancia del contexto en esta teoría no se ha reflejado en el proceso de evaluación. La *Modificación del Entorno* (Baker y Feuerstein, 1991) hace alusión, por una parte, a las expectativas de padres y educadores sobre las posibilidades educativas y de logro de un niño y, por otra, a la necesidad de implementar en caso necesario procesos de intervención específicos que aseguren adecuadas experiencias de aprendizaje mediado. Feuerstein ha manifestado públicamente que su interés se centraba en modificar las estructuras cognitivas del sujeto afectado por carencias de EAM, y que el contexto, a efectos de actuar sobre él, aunque importante, quedaba en un segundo plano (Feuerstein, 1998b). En realidad, la actuación teórica sobre el contexto desde la teoría de la EAM está fuera de toda duda, ya que la propia formulación del concepto remite a experiencias en contextos de interacción. La conceptualización de la alternativa presentada como Modificación del Entorno no parece añadir esencialmente nada a la teoría ya desarrollada. El problema, bajo nuestro punto de vista, es la no inclusión del contexto como variable en el proceso de evaluación dinámica, ya que, tal como propone Saxe (1977, 1981): *Una comprensión completa de la zona de desarrollo próximo es solamente posible si consideramos el contexto histórico específico*. En este sentido, una evaluación del proceso de enseñanza y aprendizaje que no contempla el contexto (contexto de interacción y de contenidos que conforman la realidad donde tienen lugar las acciones educativas) difícilmente puede orientar su mejora.

En definitiva, en línea con lo expuesto desde diversas fuentes (Mora, 1998; Saldaña, 2001; Sternberg y Grigorenko, 2003), la evaluación dinámica del potencial de aprendizaje y las ideas en las que se sustenta, constituyen un modelo que ha realizado importantes aportaciones al campo de la evaluación dinámica y ha facilitado y contribuido al desarrollo de nuevas pruebas de evaluación. Sin duda, Feuerstein ha contribuido de manera fundamental al hecho de que la evaluación dinámica y los principios teóricos y metodológicos que subyacen a su aplicación, sean hoy conocidos e investigados por un cada vez mayor número de psicólogos y educadores en todo el mundo.

Dentro del paradigma de la EAM se han ido desarrollando otros modelos de evaluación dinámica, de los que a su vez se han derivado también nuevas pruebas del potencial de aprendizaje que, o bien han tratado de cubrir segmentos de edad que no quedaban bajo el paraguas del LPAD, o se han centrado más específicamente en determinadas tareas

para establecer la dinámica del cambio. Estas pruebas han sido construidas por colaboradores de Feuerstein con una base sólida en la teoría de la EAM y la modificabilidad estructural cognitiva desarrollada por éste y a modo de ampliación de su campo de aplicación. Entre estas pruebas cabe destacar el Dispositivo de Evaluación del Aprendizaje en Preescolar (PLAD) de Lidz y Thomas, la Batería de Modificabilidad Cognitiva (CMB) de Tzuriel, y el modelo MindLadder de *constructivismo mediado*, de Jensen. A continuación examinaremos estos modelos que, en definitiva, han contribuido a desarrollar y perfeccionar la propuesta de Feuerstein. Dentro de esta exposición de los modelos que se han desarrollado al amparo de la teoría de la EAM, merecerá una especial atención para nosotros el modelo de *constructivismo mediado* desarrollado por Mogens R. Jensen, al que dedicamos un apartado diferenciado, en función de la mayor sintonía que encontramos con nuestra propia propuesta de evaluación dinámica.

El dispositivo de evaluación del aprendizaje en preescolar (PLAD) de C.S. Lidz y C. Thomas

Bases teóricas y metodológicas

El PLAD (Lidz y Thomas, 1987) es un instrumento desarrollado a raíz del dispositivo de evaluación del potencial de aprendizaje (LPAD) y basado por tanto en la teoría y el enfoque metodológico de Feuerstein. Además de esta notable influencia, el instrumento desarrollado hace acopio de las propuestas metodológicas presentadas por Luria (1984) y desarrolladas en el modelo PASS (Planning-Attention-Simultaneous-Successive) de Das y colaboradores (Das et al., 1994; Naglieri y Das, 1988). En este modelo se concibe el funcionamiento cognitivo en correlación con sus bases neurofisiológicas, distinguiendo tres fases: una primera fase de *atención-activación*, que tendría un sustrato neurológico localizado en el tronco del encéfalo, el diencefalo y las zonas mediales de la corteza, y que se caracteriza por permitir el input (la entrada de información) en las mejores condiciones posibles para su codificación y elaboración; una segunda fase de *procesamiento simultáneo y sucesivo*, caracterizada por la recepción, codificación y elaboración de la información, y que se correspondería neurológicamente con zonas localizadas en los lóbulos occipital, temporal y parietal del encéfalo; y una tercera fase denominada de *planificación*, que encontraría localización neurológica en los lóbulos frontales y más concretamente en la corteza prefrontal, caracterizándose por la regulación y el control de la conducta, incluyendo acciones relativas a la planificación, la supervisión y la evaluación de la propia conducta.

El PLAD está pensado para su aplicación en la educación infantil, en edades comprendidas entre los tres y los cinco años. El formato utilizado en el proceso de evaluación es también el de pretest/ mediación /posttest, aunque a diferencia del modelo de Feuerstein, se han empleado tareas de evaluación y poblaciones diana distintas para su aplicación (Sternberg y Grigorenko, 2003).

Investigación y revisión crítica del modelo

En un estudio realizado por Lidz (1991) se aplicó el PLAD a un grupo de niños y niñas del programa de educación compensatoria Head Start. La comparación tuvo lugar con otro grupo equivalente que funcionó como GC y al que se aplicaron las pruebas pretest y posttest, pero que no recibieron la intervención que forma parte de la fase de mediación del instrumento. Los resultados mostraron la mejora de los chicos del GE con

respecto al GC. Así mismo, en otra investigación realizada por Reinharth (1989) se aplicó el PLAD a un grupo de niños con retraso en su desarrollo, que fue comparado con un GC equivalente; de la misma forma que en el estudio anterior, los niños del GE experimentaron mejoras notables con respecto al GC, mejoras que fueron confirmadas una semana más tarde mediante la administración de una prueba de seguimiento. Algunos investigadores han analizado críticamente estos estudios en términos muy parecidos a los que se recogían anteriormente para el modelo del LPAD de Feuerstein (Sternberg y Grigorenko, 2003). Principalmente se destaca que en los estudios llevados a cabo por Lidz y sus colaboradores para aplicar el PLAD no se contempló ni analizó la presencia de un tratamiento alternativo para el GC que, además de neutralizar el efecto de la práctica, pudiera ofrecer unas conclusiones más potentes sobre el efecto de la fase de intervención del instrumento experimentado.

La Batería de Modificabilidad Cognitiva (CMB) de D. Tzuriel

Bases teóricas y metodológicas

La Batería de Modificabilidad Cognitiva (Tzuriel, 1995b, 2000) se ha desarrollado especialmente para ser aplicada con niños y niñas de educación infantil y lo que sería nuestro equivalente al primer ciclo de educación primaria, aunque puede emplearse con alumnado de mayor edad que presente dificultades en su aprendizaje. Este instrumento puede ser utilizado en procesos de evaluación y también de intervención (Lauchlan y Elliott, 1997), aspecto que lo hace especialmente interesante para nosotros, ya que guarda estrecha relación con nuestra propuesta. La identificación temprana de unas funciones cognitivas deficientes específicas (Feuerstein, Rand y Hoffman, 1979), que como veíamos anteriormente cuando exponíamos las líneas teóricas del modelo de Feuerstein, podrían condicionar las posibilidades de exposición a experiencias de aprendizaje mediado, así como de aprovechamiento de las mismas, es de suma importancia para promover el desarrollo de estrategias de aprendizaje y optimizar de esta forma la preparación para los aprendizajes escolares.

La CMB está compuesta por seis subtests que evalúan diferentes funciones cognitivas: (a) seriación, (b) reproducción de patrones, (c) analogías, (d) secuencias-Nivel I, (e) secuencias-Nivel II, y (f) memoria; y consta de una serie de materiales de tipo manipulativo a través de los cuales se plantean los distintos problemas a resolver. Se trata de cuatro placas de 18 cm. x 18 cm. con 9 huecos o «ventanas» cada una y 64 bloques de madera que varían en color (4 colores distintos) y en tamaño (4 tamaños). Los bloques son colocados en las «ventanas» de las placas en función del problema propuesto en los diferentes subtests.

Investigación y revisión crítica del modelo

El CMB ha sido utilizado en varios estudios para evaluar el potencial de aprendizaje y, como apuntábamos anteriormente, también se ha investigado su validez como instrumento de evaluación e intervención (Jeffrey, 1997; Lauchlan y Elliott, 1997; Jeffrey y Tzuriel, 1999; Tzuriel, 1997). Lauchlan y Elliott (1997) realizaron un estudio comparando dos grupos experimentales, que recibieron una intervención con programas de educación cognitiva, con un GC que recibió una enseñanza de tipo tradicional. De los dos GE, el primero (N=15), con edades de 8 y 9 años, recibió una intervención basada en el CMB y el segundo GE (N=15) recibió el *Programa de Enriquecimiento*

Instrumental (PEI) de Feuerstein, Rand, Hoffman y Miller (1980). El GC estaba compuesto por 30 alumnos. Todos los sujetos de la muestra asistían a clases de apoyo para alumnado con dificultades de aprendizaje. Los chicos fueron evaluados inicialmente con otro instrumento creado por el propio Tzuriel, el CATM (Children's Analogical Thinking Modifiability) (Tzuriel y Klein, 1985, 1987) y asignados posteriormente a los grupos control y experimental, controlando la equivalencia en función de las puntuaciones obtenidas. Tras la fase de tratamiento, los resultados obtenidos en la evaluación dinámica posterior mostraron que el CMB era un instrumento capaz de producir cambios cognitivos significativos en el grupo al que fue aplicado. Es de suponer que la aplicación del PEI produciría también resultados superiores en el GE al que se aplicó, que los obtenidos por el GC, pero no disponemos de estos datos, así como tampoco de los que permitirían comparar los tratamientos CMB y PEI.

En un estudio llevado a cabo por el propio Tzuriel (1999) con niños y niñas de educación infantil y primer curso de educación primaria, se trataron de establecer las propiedades psicométricas de los subtests del CMB, así como las relaciones que éstos mantenían entre sí en la explicación de los resultados; se exploró la magnitud de los cambios producidos entre las fases pre y postest para cada uno de los subtests de la batería y, por último, se pretendía estudiar la validez predictiva de la prueba con respecto al rendimiento en comprensión lectora y matemáticas en el primer curso de primaria. La muestra, heterogénea en cuanto al nivel socioeconómico de las familias, estaba compuesta por 92 chicos y chicas de infantil, pertenecientes a 26 clases distintas, y 131 de primaria, representando un total de 28 clases de primero. Se controló el nivel educativo y socioeconómico de las familias. El CMB fue aplicado individualmente por evaluadores participantes en un seminario sobre ED. Todos los alumnos recibieron los subtests del CMB en el mismo orden, constituyendo el total de tiempo empleado en la administración de la prueba unas cinco o seis horas para cada alumno. Los resultados concernientes al análisis de la magnitud del cambio entre las fases del estudio para cada uno de los subtests mostraron en general unas intercorrelaciones de bajo nivel, reflejando una débil relación entre los índices de modificabilidad de cada uno de los subtests. Estos resultados apuntarían en la dirección de una especificidad de dominio en la modificabilidad cognitiva. Por otra parte, en cuanto al análisis de la validez predictiva del CMB para el rendimiento en comprensión lectora y en matemáticas, fue administrado a una submuestra de alumnos de primero de primaria una prueba estandarizada de comprensión (Ortar y Ben-Schachar, 1972) y otra de matemáticas (Kidron, 1989). Las variables predictoras fueron las puntuaciones pretest y postest del CMB. Los resultados mostraron que las puntuaciones postest en los subtests de *seriación* y *analogías* predecían el rendimiento en comprensión lectora y que las puntuaciones postest de *seriación* y las pretest de *reproducción de patrones* predecían el rendimiento en matemáticas. El porcentaje de varianza explicada por el CMB para la comprensión lectora fue del 45% y del 57% para matemáticas.

En definitiva, que duda cabe sobre la decidida intención de los colaboradores y seguidores de Feuerstein de mejorar las propiedades psicométricas y de aplicación de las pruebas que han construido a raíz de la teoría de la EAM. A pesar de esto, y de acuerdo con la reflexión del propio Tzuriel: *la principal validación de la ED debería focalizarse sobre la predicción de aspectos que no son predichos por los resultados de las pruebas estáticas* (Tzuriel, 2000; pág. 404). Relacionado con esto, podría pensarse que, a pesar de los logros conseguidos, todas estas pruebas que han sido desarrolladas al

abrigo del LPAD adolecen de un mismo factor: la contextualización de los instrumentos de evaluación y, por tanto, de mayores posibilidades de establecer criterios de generalización de sus resultados y de validez utilizando precisamente criterios de rendimiento escolar y solución de problemas reales.

Un apunte que nos parece de especial importancia con respecto a las conclusiones que se exponen en el último estudio presentado en esta revisión: Como es sabido, es necesario tener cierta cautela con la interpretación que hacemos de datos correlacionales como los que se aportan para establecer la validez predictiva del CMB con respecto a dominios como la comprensión lectora o las matemáticas. A parte del elemento ya comentado de la descontextualización de las tareas que sirven para la ED en este instrumento, el hecho de que el contenido de las actividades que componen la batería — recordemos: bloques de madera de colores y tamaños distintos—, no guarde en principio relación con la comprensión de textos, podría al menos hacernos dudar de la validez que aportan en realidad las correlaciones presentadas; por ejemplo, cuando con base en las correlaciones halladas entre las dificultades perceptivo-visuales o grafo-motrices y los problemas de lectura, se llegaban a realizar interpretaciones de tipo causal, la lógica de estas inferencias llevaba a plantear tareas basadas en el entrenamiento de habilidades de tipo perceptivo y motriz para la mejora de la lectura (Alegría, 1985; Gómez, Valero, Buades, y Pérez, 1995). Sólo cuando se acumularon las evidencias que minimizaban la incidencia de las dificultades anteriores en la explicación de los problemas de lectura a favor de otras que, mediante la aportación de numerosos datos de estudios experimentales, apuntaban claramente a la existencia de dificultades de tipo psicolingüístico y, por tanto, más directamente relacionados con los elementos que forman parte integrante de la lectura (fonemas, palabras, etc.), se comenzó a dar un giro a las prácticas educativas y reeducativas para mejorar los procesos de lectura y comprensión (Jiménez y Ortiz, 1998). En este caso, dado el carácter experimental de los estudios realizados, sí se cumplía que el entrenamiento en conciencia fonológica, derivado de las conclusiones sobre los procesos relevantes que subyacían a las dificultades de lectura, se relacionaba directamente con la mejora en pruebas de comprensión lectora. Por tanto, la intervención apuntada desde un proceso de evaluación que conecta causalmente dificultades en la conciencia fonológica y problemas en la lectura, no puede ser otra que incidir en la mejora de los procesos de conciencia fonológica.

Siguiendo la lógica expuesta anteriormente, en el caso del estudio de Tzuriel, en el que los subtests de seriación y analogías correlacionaban con el rendimiento en la prueba criterio de comprensión lectora, ¿plantearíamos un entrenamiento en actividades de seriación o en analogías, realizadas con bloques de madera, para mejorar la comprensión lectora? Pensamos que difícilmente adoptaríamos esta opción. La validez predictiva de una prueba de evaluación con respecto a un dominio específico y complejo como la lectura, no puede estar basada exclusivamente, a nuestro juicio, en tareas no contextualizadas y que no guardan relación con los elementos o procesos que conforman ésta; incluso en el caso en que se demostrara que los procesos cognitivos implicados en la realización de las seriaciones o de las analogías, también lo están en la comprensión de textos, esto difícilmente explicaría su influencia en el proceso de comprensión, debido a la especificidad comentada anteriormente, y máxime cuando esos procesos cognitivos empleados para resolver las tareas de seriación y analogías han operado sobre elementos ajenos al lenguaje escrito. El establecimiento de validez predictiva sobre la comprensión de textos debería contar a nuestro parecer con

actividades de ED contextualizadas, y, por tanto, basadas en contenidos que se relacionen significativamente con su uso en contextos de enseñanza y aprendizaje reales, ya sean escolares o sociales; así mismo, deberían ser actividades cuyo contenido esté directamente relacionadas con el lenguaje escrito o con los elementos y procesos que lo conforman.

El modelo de evaluación dinámica MindLadder, de M.R. Jensen

Bases teóricas y metodológicas

El modelo MindLadder (Jensen, 1992; 1998; 2000) parte de la teoría de la EAM para proponer la noción de *constructivismo mediado*. Con este concepto, el autor quiere situar el énfasis en la mediación del aprendizaje de la capacidad de construir nuevos aprendizajes y de usar con una mayor eficacia el propio conocimiento. Se plantea para esto el objetivo de conocer y explorar el funcionamiento cognitivo de un sujeto y cómo éste puede llegar a aprender a construir y usar más eficazmente sus conocimientos. El modelo se basa en cinco de las características de la EAM para desarrollar la construcción del conocimiento: (1) *Intencionalidad-Reciprocidad*, a través de la cual se forma un vínculo entre el mediador y el alumno; (2) *Trascendencia*, por medio de la cual el alumno desarrolla un nuevo modo de funcionamiento cognitivo que va más allá de la tarea o la situación de aprendizaje concreta; (3) *Mediación del significado*, que crea en el alumno la necesidad de desarrollar un nuevo modo de funcionamiento cognitivo; (4) *Mediación del sentido de competencia*, a través de la cual se ayuda al alumno a afrontar nuevas situaciones de aprendizaje; y (5) *Regulación y control mediado de la conducta*, por la que tiene lugar la instauración de una nueva forma de pensar y actuar que amplía el horizonte cognitivo.

Ha sido desarrollado para ser empleado con población escolar con y sin necesidades educativas especiales y, según su autor, su diseño responde a la necesidad de establecer de forma integral una evaluación del desarrollo actual de las capacidades de aprendizaje del alumnado, de la adquisición del conocimiento y de las habilidades de resolución de problemas. Está pensado para clarificar y ayudar en el proceso de instrucción, identificando las funciones de construcción del conocimiento que se hallan en proceso de desarrollo y, por tanto, requieren mediación. No fue en cambio pensado como instrumento que permitiera clasificar mejor al alumnado, ni para valorar el logro de un alumno con relación a un determinado grupo (Jensen, 2000). Este punto nos resulta de especial interés y volveremos sobre él más tarde, cuando abordemos nuestra propuesta de evaluación dinámica de procesos implicados en la lectura.

Por otra parte, Jensen propone que una vez establecidas y delineadas en cierta forma las funciones cognitivas, e investigada la forma en la que éstas pueden desarrollarse mejor, los educadores pueden crear o poner en práctica un currículum escolar que garantice mayores y más eficaces oportunidades de aprendizaje, es decir, en este caso la evaluación dinámica quiere establecer una conexión más clara y directa con el rendimiento y la educación escolar, de la que podíamos desprender del análisis de la teoría y el interés investigador de los estudios de Feuerstein. En este sentido, la propuesta realizada, según su propio autor, puede ser tanto un instrumento de evaluación como una base para el diseño y la implementación de pruebas de aprendizaje incardinadas en la escuela.

Como apuntábamos al comienzo, la teoría desarrollada acerca del constructivismo mediado (Jensen, 1992; 1998) enfatiza la necesidad de establecer mediación en el desarrollo de los instrumentos de construcción de las funciones cognitivas del alumnado. Tres son los tipos de funciones cognitivas que se han descrito desde esta teoría: (a) *funciones intelectivas*, que incluirían a su vez las funciones de recepción, como la atención o la conducta exploratoria, las funciones de transformación, como la representación mental y la planificación, y las funciones de comunicación, como los instrumentos verbales y la autorregulación; (b) *funciones no intelectivas*, como las motivaciones, las necesidades y los atributos de personalidad; y (c) *funciones de actuación*, como la rapidez, la precisión y la formación de hábitos (Sternberg y Grigorenko, 2003).

El constructivismo sugiere que *el conocimiento y el aprendizaje son, en buena medida, el resultado de una dinámica en la que las aportaciones del sujeto al acto de conocer y aprender juega un papel decisivo* (Coll, 2001). La teoría del constructivismo mediado, como decíamos anteriormente, hunde sus raíces en la teoría de la EAM, considerando que los seres humanos construyen experiencias y regulan progresivamente sus acciones usando sistemas de signos y símbolos que toman significado con base en la cultura (Jensen, 1992). La teoría trata de explicar el proceso por el cual la EAM contribuye al desarrollo de las funciones de construcción del conocimiento mediante las cuales el sujeto puede procesar la información, beneficiarse de la experiencia, aprender a aprender y aprender a modificar sus propias estrategias de aprendizaje. La EAM podría jugar, de esta forma, un papel fundamental para asegurar la flexibilidad de los esquemas que, a su vez, permiten los procesos de asimilación y acomodación descritos por Piaget (Jensen, Feuerstein, Rand, Kaniel y Tzurriel, 1988).

Por otra parte, la teoría se propone describir y acometer cómo las funciones cognitivas mediadas llegan a ser integradas y usadas con progresiva eficacia dentro de la estructura de conocimientos del sujeto. En este sentido, como comentábamos anteriormente, la teoría del constructivismo mediado propone que la evaluación dinámica es parte integrante del proceso de enseñanza y aprendizaje en el aula. Junto con los procesos instruccionales, la evaluación dinámica de Jensen comparte el objetivo de desarrollar mayores niveles de rendimiento académico en cada una de las áreas mediante estrategias de construcción del conocimiento que contribuyen a un aprendizaje eficaz y duradero. De esta forma, se propone un currículum escolar basado en los principios del constructivismo y del aprendizaje mediado. Los educadores usan el currículum como base para ayudar a su alumnado a adquirir progresivamente un mayor control del proceso de aprendizaje, una actitud de búsqueda activa de significados y un mayor dominio de la comprensión y el uso del conocimiento. El modelo de currículum y la modalidad de aprendizaje propuestos tratan de proporcionar al alumnado oportunidades para la adquisición de una amplia base de conocimientos de tipo conceptual y también experiencial. Dentro de este contexto, el educador proporciona EAM para ayudar a sus alumnos a desarrollar funciones de construcción de conocimientos mediante la puesta en práctica de cada una de las características del aprendizaje mediado expuestas anteriormente. Por su parte, en el proceso de evaluación dinámica, Jensen propone un contexto similar, en el que se trata de facilitar el desarrollo de funciones constructivas del conocimiento utilizadas para organizar y ensamblar el conocimiento declarativo y *convertirlas* mediante la experiencia proporcionada en funciones de conocimiento procedimental.

Una de las innovaciones a todas luces más interesantes en el campo de la ED es la utilización del ordenador en la recogida y el análisis de los datos del proceso acumulados durante la aplicación de las actividades. El programa informático CAMET (*computer Assisted Modifiability Enhancement Techniques*) es definido por su autor como un programa dinámico e interactivo que es utilizado en un entorno de aprendizaje y evaluación con el alumno, y que permite registrar y analizar la transición del sujeto en el dominio de la tarea a través de sus respuestas, empleando para ello pruebas paramétricas. En el desarrollo de la actividad dinámica propuesta a través de este programa, el evaluador se sitúa junto al alumno compartiendo y estableciendo una determinada mediación ante una tarea a realizar; de esta forma, el uso de este programa no está diseñado para suplantar al evaluador, ya que se parte de una concepción en la que la interrelación entre éste y el alumno es esencial; sin embargo, según su autor, su uso permitiría mejorar la información y optimizar el aprovechamiento que alumnos y evaluadores pueden hacer de la ED.

En cuanto a las tareas propuestas, el programa contiene tres niveles de dificultad; éstos están pensados para adaptarse gradualmente a distintos niveles de necesidades educativas relacionadas con la construcción del conocimiento. En el cuadro 12 hemos recogido los instrumentos que forman parte del modelo MindLadder propuesto por Jensen.

Investigación y revisión crítica del modelo

El constructivismo mediado surge de la certeza de que las funciones cognitivas del alumnado podrían ser desarrolladas y modificadas mediante el uso de materiales e instrumentos específicos (Feuerstein y Jensen, 1980); sin embargo, esta premisa fue revisada a la luz de un estudio longitudinal llevado a cabo por la Universidad de Yale (Jensen y Singer, 1987; Jensen, 1990). El estudio fue realizado con 53 alumnos de educación especial que durante tres años habían estado recibiendo una intervención mediante un programa para la mejora de sus habilidades cognitivas y de aprendizaje. Los resultados apuntaban a la necesidad de establecer una mayor conexión entre el desarrollo de las funciones cognitivas y el desarrollo del conocimiento y las estrategias. Se estableció una comparación entre aquellos alumnos del GE del estudio (N=26) que habían recibido una mayor cantidad de intervención, y aquellos que recibieron una menor cantidad (N=27), indicando los resultados, tanto la valoración de los profesores como las pruebas aplicadas, diferencias muy significativas a favor del primer subgrupo en cuanto a competencia intelectual, disposición positiva hacia el logro, funcionamiento autónomo y otras funciones cognitivas; sin embargo, cuando los resultados fueron comparados con el GC (N=50) no se encontraron, con respecto a éste, evidencias significativas de mejora en ninguno de los dos subgrupos para ningún parámetro académico analizado. Las medidas obtenidas por el GE correlacionaban con otras medidas derivadas de la propia aplicación del programa, así como con medidas obtenidas usando tareas similares, lo que supone en palabras del autor y usando la terminología de Campione y Brown (1987), una transferencia «cercana» de los aprendizajes; pero en cambio, no correlacionaban con medidas derivadas de actividades académicas, lo que, en tal caso, hubiera supuesto una transferencia «lejana» del aprendizaje. La clarificadora interpretación que realizan de todos estos resultados los autores del estudio y, posteriormente, el propio Jensen (2000), hace referencia a la dificultad para establecer puentes entre los contenidos del programa y el currículum escolar, de forma que el aprendizaje obtenido no se mostraba útil ni significativo para su

Cuadro 12. Instrumentos de evaluación dinámica en el modelo MindLadder. Tomado de Jensen (2000)

Nombre del instrumento	Descripción	Opciones y características
* Matrices	* Consiste en la presentación de analogías no verbales con un formato de 4 ó 9 elementos. El rango de dificultad de los ítems va desde el tipo <i>Gestalt</i> hasta los que implican un mayor razonamiento inductivo.	* Las opciones de R (se dispone de 6 a 8) pueden mostrarse tapadas o no. Los ítems seleccionados pueden ser mostrados en la matriz o indicados numéricamente.
* Plantillas	* Consiste en la superposición de diferentes plantillas de colores con diferentes recortes para recrear un modelo.	* Los alumnos pueden separar las plantillas o volver a empezar. Las plantillas seleccionadas pueden ser mostradas o anotadas en la hoja de R para ser revisadas posteriormente.
* Figuras realizadas con puntos	* Consiste en conectar puntos para formar figuras geométricas mostradas en un modelo. La complejidad de la tarea es progresiva en cuanto al número y organización de los puntos para formar las figuras.	* Es posible usar el ratón, dibujar de punto a punto o saltar con el dedo de un punto a otro. Las líneas y los puntos pueden ser organizados en orden inverso.
* Tablero lógico	* Consiste en la descripción de objetos comunes y la relación que guardan entre ellos. Se necesitan dos o más premisas para reducir o eliminar posibilidades de error y así fijar la localización de cada objeto. Las tareas varían de 3 a 9 objetos y emplean una variedad de expresiones lógicas y sus negativas para describir las relaciones entre ellos.	* Las premisas pueden ser presentadas todas a la vez o una a una en la pantalla. A su vez, de debajo de cada una puede ser extraída información.
* Lector	* Consiste en la presentación de uno o más párrafos describiendo acontecimientos cotidianos. Existen 8 niveles de dificultad y, tanto las palabras como las estructuras gramaticales usadas en las historias presentadas, son utilizadas en la formación de los ítems de las diferentes partes de la evaluación. Hay un total de 40 ítems.	* Las palabras en el texto pueden ser marcadas y mostradas por separado. Como se realiza una lectura en voz alta, este instrumento no permite el registro informático de los datos.
* Sustituciones	* Consiste en presentar una serie de viñetas con líneas o iconos de objetos comunes. Primeramente los ítems son asociados con sus sustitutos uno a uno; posteriormente, los alumnos emparejan ítems y sustitutos recordando y estableciendo sus conexiones con éstos presentes, y por último intentan recordar y establecer las conexiones sin la presencia de los objetos.	* La muestra de objetos y sustitutos puede ser controlada; así mismo, la duración de la muestra y las pausas entre éstas puede ser registrada. Los alumnos pueden ojear los sustitutos no mostrados, así como examinar previamente los iconos mostrados.
* Copia	* Consiste en recomponer los puntos y líneas que son presentados en un modelo. Las tareas varían en dificultad en función del número de puntos y líneas a conectar, así como en el tamaño de los puntos utilizados.	* Los alumnos pueden usar los dedos o el ratón para trazar y conectar las líneas; así mismo, éstas y los puntos pueden ser organizados en orden inverso.
* Gráfico	* Consiste en trazar coordenadas y líneas dentro de un diagrama cartesiano. Las tareas varían desde una a diez coordenadas y más de quince líneas conectadas entre ellas.	* Los alumnos pueden usar los dedos o el ratón para trazar las líneas y coordenadas. Líneas y coordenadas pueden ser organizadas en orden inverso.

aplicación en nuevos contextos de aprendizaje. El estudio sugiere por tanto la necesidad de acercar o, más allá, incluir el desarrollo de la cognición en el contexto del desarrollo del conocimiento y la aplicación de ese conocimiento. El constructivismo mediado tratará a partir de entonces de acometer esta empresa.

Entre los estudios que han tenido lugar para desarrollar este cometido, algunos han utilizado precisamente los cursos escolares como medida criterio. En una investigación realizada en California con 25 niños se evaluó su competencia en lectura y matemáticas mediante el SAT (Stanford Achievement Test) al final del segundo, tercer y cuarto curso (Doerr y Jensen, 1998). Al final de segundo el grupo fue dividido en dos (Grupo A; N=13 y Grupo B; N=12) y se aplicó el test indicando diferencias significativas a favor de uno de los grupos (A); posteriormente, este grupo, durante el tercer curso, trabajó con un programa experimental mediante la aplicación de la teoría del constructivismo mediado, mientras que el Grupo B trabajó en otra clase de tercero con otros profesores mediante una enseñanza más tradicional. Al finalizar tercero las diferencias, tanto en lectura como en matemáticas habían aumentado entre los dos grupos. Al iniciar cuarto, los grupos fueron reunidos y recibieron clase con el profesor que había estado utilizando el programa basado en el constructivismo mediado ya desde tercero con el Grupo A. Pasado el cuarto curso se volvió a evaluar y esta vez no se hallaron diferencias entre ambos grupos ni en lectura ni en el área de matemáticas. Los resultados no se debieron a un menor desempeño de los alumnos integrantes del Grupo A, ya que éstos mantuvieron el nivel alto que, según los indicadores del SAT, habían estado obteniendo en las anteriores valoraciones, y muestran que en un año de aplicación del programa, los alumnos que hasta entonces habían obtenido puntuaciones muy por debajo de sus compañeros, lograron alcanzarles, incrementando sus puntuaciones de forma muy significativa.

A pesar de que nuevos estudios deberán ser realizados y validados, probablemente con un mayor número de sujetos y con diversos grupos de alumnos que presenten distinto nivel de competencia curricular; y a pesar de que, así mismo, tanto la taxonomía utilizada por Jensen para diferenciar las funciones cognitivas, como su contenido concreto, no parecen apoyarse en estudios empíricos sólidos que establezcan su validez (Sternberg y Grigorenko, 2003); lo cierto es que este modelo realiza, a nuestro juicio, importantes aportaciones al campo de la ED, sobre todo a nivel de conceptualización. En primer lugar, trata de establecer una conexión entre la teoría de la EAM y el constructivismo; para ello, como hemos visto, desarrolla un modelo de constructivismo mediado que enfatiza y facilita la participación del alumno en actividades de aprendizaje significativo junto con la mediación de las funciones cognitivas que permiten la construcción del conocimiento. En este sentido, el modelo propuesto por Jensen encaja formidablemente en la conceptualización de *constructivismo dialéctico*, *constructivismo social* o *socioconstructivismo* (Newman, Griffin y Cole, 1989; Coll y Solè, 1993; Coll, 2001; Solè y Teberosky, 2001; Goodenow, 1992; Salomon y Perkins, 1998; Luque, Ortega y Cubero, 1997; Cubero y Luque, 2001; Bruning, Schraw y Ronning, 2002; Navarro, 2004), formulada a raíz de los trabajos de Vygotski. Ya hacíamos mención en la parte introductoria de este capítulo a la concepción que sobre el desarrollo y el aprendizaje de las funciones cognitivas se fue elaborando por Vygotski y sus más íntimos colaboradores, así como por los investigadores que, seducidos por su obra, han contribuido posteriormente a ampliar significativamente los horizontes de ésta. El constructivismo dialéctico, en sintonía con la teoría de Vygotski, sitúa el desarrollo de las funciones cognitivas superiores en la interacción que mantienen

aspectos internos y externos del sujeto; de esta forma, las estructuras relacionales de intercambio comunicativo y de aprendizaje mediado que tienen lugar entre miembros de una determinada cultura, pasan a formar parte de las estructuras cognitivas del individuo a través de un proceso de interiorización que éste realiza gracias a la mediación semiótica que ejercen los miembros expertos de la cultura, situándose en esa *región dinámica sensible al aprendizaje de aptitudes culturales*, que es la ZDP o, de otra forma llamada, la «zona de construcción».

El proceso de transformación cognitiva tendría lugar gracias a esta interacción mediada que se da entre miembros expertos y novatos de la cultura. La estructura y los contenidos de la interacción se interiorizan y pasan a integrar la estructura de procesos y conocimientos del niño. Por su parte, el proceso de interiorización del conocimiento de la cultura a partir de esta interacción mediada fue descrito por Leontiev, uno de los primeros colegas de Vygotski. Leontiev (1981), propuso el término *apropiación* en sus trabajos para hacer referencia al proceso que el niño efectúa sobre los contenidos culturales que han sido generados históricamente por la sociedad. Este término de apropiación quiere expresar que el proceso de interiorización del conocimiento cultural en la ZDP no es un traspaso literal de las estructuras externas a la cognición del individuo. Por el contrario, se hace hincapié en que el niño realiza su propia comprensión de los acontecimientos y de los contenidos que forman parte de las interacciones sociales. De esta forma, los niños pueden participar activamente en situaciones de interacción que no captan ni comprenden en su totalidad (p.e. situaciones en las que una madre muestra y lee un cuento a su hija), pero que de forma progresiva influyen en ellos. Los niños entran en contacto, a través de esta participación activa en situaciones de intercambio comunicativo, con formas de pensamiento más avanzadas, que enriquecen y posibilitan el cambio de sus estructuras cognitivas.

En segundo lugar, esta teoría del constructivismo mediado se ha interesado por la contextualización de las intervenciones dirigidas a la mejora de las funciones cognitivas, lo que en línea con lo propuesto en nuestro propio trabajo, parece ser de suma importancia si se quiere conseguir una cierta generalización de los resultados obtenidos, tanto en lo referente a procesos de evaluación como de intervención. En este sentido, la propuesta de Jensen sobre el proceso de ED contempla la colaboración directa del profesorado de los alumnos que son objeto de valoración; de este modo, se solicita del profesor en primera instancia la cumplimentación de un cuestionario (*Primary Sources Inventory record - PSI*), en el que éste debe designar cuáles son a su juicio las funciones cognitivas y de aprendizaje que son deficitarias en el alumno. Posteriormente, el evaluador coteja y discute junto al profesor sobre las dificultades del alumno, que después son valoradas con el CAMET. Los resultados obtenidos son nuevamente valorados con el profesor y de este trabajo se derivan una serie de orientaciones psicopedagógicas para lograr optimizar el desarrollo cognitivo del niño. En un estudio de caso presentado por Jensen (2000) se recoge este proceso con un niño de doce años con especiales problemas de lectura según su profesor. Los resultados del estudio mostraron que, tras la fase de intervención con el CAMET, en la que se incluyó la actividad *Reader* (lector) para la valoración dinámica de una tarea de comprensión, y poder, según el autor, obtener información contextualizada en el dominio en el que presentaba especiales dificultades este alumno, más allá de los datos ofrecidos por la evaluación mediante la utilización de tareas sin relación aparente con la lectura; pudieron ser ofrecidas al profesor indicaciones metodológicas para lograr mantener en el tiempo los logros significativos que durante la valoración dinámica había obtenido el

chico, tanto en las tareas descontextualizadas como en la de lectura. El autor, aunque consciente de que las condiciones de valoración dinámica individualizada en las que se había desarrollado el proceso, no son equiparables a las que caracterizan el trabajo con un grupo clase, apunta que los datos obtenidos pueden guiar la intervención y el apoyo que el alumno necesita, y revelan lo que cabe esperar del alumno con una intervención similar a la utilizada. En definitiva, de esta forma la contextualización de las intervenciones propuestas por Jensen en su modelo de ED, proporcionaría una mayor validez ecológica a los resultados.

En tercer lugar, se propone un programa de intervención que trata de establecer conexiones con el currículum escolar y cuyos contenidos parten del análisis de las funciones cognitivas que se quieren desarrollar y que, a su vez, pueden ser valoradas en el proceso de ED. Por último, en cuarto lugar, la ED según este modelo es parte integral del proceso de enseñanza y aprendizaje y persigue finalmente los mismos objetivos que la implementación del programa de intervención que se ha derivado de su propuesta, es decir, la mejora de las funciones de construcción del conocimiento.

El modelo de evaluación dinámica basada en el currículum (ACFS), de Carol S. Lidz

Bases teóricas y metodológicas

Uno de los desarrollos de la ED que más interés ha despertado en los últimos años ha sido el de las pruebas de evaluación elaboradas a raíz del análisis del currículum escolar y, por tanto, de aquellos aprendizajes que los niños deben realizar en la escuela y por los que serán valorados sus logros y sus conocimientos. Este tipo de evaluación que refleja los objetivos educativos del currículum, parece encontrarse más cerca de la realidad escolar y podría constituirse en una especie de *eslabón perdido* que, una vez hallado y pulido, podría contribuir de manera significativa a enlazar los procesos de evaluación e intervención (Elliot y Fuchs, 1997; Fuchs y Fuchs, 1996). Dentro de este tipo de pruebas, Lidz y Jepsen (1997) han desarrollado la *Aplicación de la Escala de Funciones Cognitivas (ACFS)*. La elaboración de esta prueba dinámica estuvo precedida del análisis del currículum escolar de la educación infantil en Estados Unidos, donde tiene su origen. El campo de aplicación de la ACFS abarca el equivalente al segundo ciclo de nuestra educación infantil (3-5 años), aunque los autores remarcan la posibilidad de utilizar la escala con alumnado mayor que presente necesidades educativas especiales.

Por otra parte, Lidz (1991, 2000), destacando la relevancia del desarrollo de este tipo de pruebas, hace alusión a un cierto riesgo que conlleva el hecho de que el contenido de las mismas esté «demasiado cerca» del currículum, confundiendo con éste y convirtiéndose de esta forma en una prueba de evaluación de los aprendizajes escolares que usan los profesores para valorar el grado de consecución de los objetivos educativos (Fuchs y Deno, 1994). Según la autora, una prueba dinámica basada en el currículum debe ir más allá, caracterizándose, no solo por otorgar información acerca de lo que el alumno conoce sobre el contenido escolar o cuál es el estado actual de las funciones cognitivas relacionadas con esos contenidos; sino de forma especial por la valoración de la capacidad y el modo de aprendizaje que muestra el alumno en tareas relacionadas con los contenidos escolares, así como de las dificultades encontradas en el camino y que impiden que tenga lugar el proceso de forma óptima. De esta forma, el desarrollo de una aproximación dinámica de evaluación que tenga en cuenta el currículum escolar y, así

mismo, los procesos que subyacen al aprendizaje de sus contenidos, sería la forma idónea de evaluación (Lidz, 2000).

El desarrollo de esta prueba recoge las influencias de los trabajos de Vygotski y Feuerstein, así como de otros investigadores que se han interesado por el desarrollo cognitivo, especialmente por el papel que juegan las funciones metacognitivas en este desarrollo. El procedimiento de evaluación proporciona la oportunidad de evaluar el desarrollo de funciones cognitivas que están relacionadas con el currículum escolar de educación infantil, así como de valorar la capacidad de aprovechamiento de una situación de aprendizaje mediado. En este sentido, nos parece importante destacar que la evaluación de las funciones cognitivas *per se* no es considerado lo esencial en esta propuesta; se trata, más bien, de valorar sobre todo la modificabilidad de estas funciones cognitivas y la calidad de respuesta del alumno a la interacción mediada. Los objetivos de la propuesta se centran en poder orientar la intervención educativa, ofreciendo información valiosa sobre cómo pone en práctica el alumno funciones cognitivas y metacognitivas en tareas relacionadas con el currículum, y sobre cómo modifica su propia actuación en función de la mediación. El proceso de evaluación seguido con la ACFS proporcionaría claves a los educadores para la creación de ZDP que permitan incrementar el desarrollo de las funciones evaluadas.

Desde este modelo de ED se asume que en el desarrollo de las funciones cognitivas del individuo tiene lugar un modo específico de procesamiento y organización de la información que viene dado por las características individuales del sujeto y que a su vez se relaciona con las experiencias y con su entorno (Lidz, 2000). Esta «construcción» personal en función de la interacción de las condiciones del propio organismo y de las experiencias de aprendizaje, hace que se acentúe especialmente la acción que es posible ejercer desde el contexto social de interacción. En definitiva, se busca que el modelo ACFS sea un modo de proporcionar, durante el curso de la fase de intervención, experiencias de aprendizaje «manipuladas» dentro de un determinado contexto social y mediante una interacción mediada.

La escala ACFS consta de seis subescalas que se corresponden con tópicos presentes en la etapa de educación infantil. Los indicadores de evaluación que ofrece la aplicación de la escala hacen referencia, por una parte al grado de aptitud que muestran los alumnos en el desarrollo de cada uno de los procesos de aprendizaje representados y, por otra, a la respuesta cualitativa del alumno a la intervención. En este sentido, los autores han descrito una serie de indicadores que permiten valorar cualitativamente la calidad de la interacción y de la respuesta del alumno a la misma en cada una de las tareas de la escala. Estos indicadores son valorados durante el pretest y en la fase de intervención. De esta forma, a lo largo del desarrollo de las tareas propuestas, el evaluador puede valorar y comparar en función de las distintas tareas, la posible puesta en práctica diferencial de estos indicadores. El contenido de las seis subescalas de la prueba ACFS, así como los indicadores cualitativos de las conductas del alumno durante la realización de las tareas, quedan recogidos en el cuadro 13.

El diseño utilizado en esta prueba contempla una fase de intervención en cada una de las subescalas, que se intercala entre un pretest y un postest que son aplicados en forma estática con instrucciones estandarizadas para su administración. El contenido de las pruebas pretest y postest es el mismo para algunos de los procesos evaluados y similar en otros. En cambio, los materiales con los que se trabaja en la fase de intervención son

diferentes a los utilizados para obtener las valoraciones inicial y final. El entrenamiento, por tanto, se efectúa sobre los procesos cognitivos, principios y estrategias que subyacen al aprendizaje de los contenidos valorados en la ACFS, pero sin utilizar materiales correspondientes a las pruebas estáticas pretest o postest. La fase de intervención y, por tanto, la interacción mediada propuesta entre el evaluador y el alumno, tiene una estructuración intermedia según los autores, de forma que, por una parte se prefijan una serie de indicaciones del proceso de intervención que facilitan la aplicación y permiten un cierto grado de estandarización del proceso y, por otra, se promueve una interacción flexible y cercana, en la que el evaluador trata de establecer una relación de empatía con el alumno, minimizando en lo posible el efecto negativo que puede llegar a causar en el niño una situación de prueba.

Cuadro 13. Subescalas y tareas de la prueba ACFS e Indicadores cualitativos de conducta. Tomado de Lidz (2000).

Tareas del ACFS

1. *Clasificación.* Al niño se le pide que clasifique unos bloques en grupos. Durante la fase de intervención, los bloques poseen características sobresalientes que ayudan a focalizar la atención del alumno sobre las reglas que debe utilizar para agrupar. Se puntúa la capacidad para deducir características distintivas, así como la flexibilidad cognitiva.
2. *Perspectiva.* Se le pide que haga de maestro y que enseñe al evaluador cómo realizaría el dibujo de un niño (que se le proporciona). Durante la intervención se proporciona un modelo de cómo comunicar algo a alguien para influir en su actuación, utilizando un dibujo de un gato. Se puntúa la capacidad para compartir el propio pensamiento con otra persona, asumiendo el rol de maestro.
3. *Memoria auditiva inmediata.* Se solicita del alumno que nos relate una pequeña historia que ha sido leída por el evaluador. Durante la intervención se enseña al niño a construir un modelo de símbolos para la historia y a usar imágenes visuales para facilitar el recuerdo. Se puntúa la capacidad de recuerdo de detalles de la historia, así como el desarrollo secuencial de la misma.
4. *Memoria visual inmediata.* Se solicita el recuerdo de 8 juguetes que se sitúan en frente del niño y posteriormente son escondidos. Durante la fase de intervención se trabajan estrategias de repetición e imágenes visuales. La puntuación incluye evidencias de conciencia de la necesidad de aplicación de estrategias de memoria, así como el número de ítems recordados.
5. *Planificación verbal.* Se solicita del alumno que relate un plan para hacer galletas (u otro tipo de comida si no está familiarizado con la propuesta). La intervención está basada en la secuenciación de los pasos necesarios para realizar un plan y en la utilización de palabras apropiadas para trazar un plan. La puntuación incluye el conocimiento del alumno acerca de distintas posibilidades en la secuencia de pasos necesaria para lograr el objetivo.
6. *Realización de secuencias de patrones.* Se solicita del alumno que complete una secuencia repetida de patrones comenzada por el evaluador, usando tangrams. Durante la fase de intervención se ayuda al niño a encontrar sentido a la secuencia y a aprender a usar las señales o claves proporcionadas por el evaluador para determinar que patrón viene después. La puntuación incluye la capacidad para justificar la respuesta.

Indicadores cualitativos de conducta

- *Autorregulación.* Se evalúa el mantenimiento de la atención durante el desarrollo de la tarea y la calidad de la interacción con los materiales (impulsividad o reflexión)
- *Persistencia.* Se evalúa una ejecución de la tarea que no busca terminarla de forma apresurada.
- *Tolerancia a la frustración.* Se evalúa la capacidad de reponerse ante la frustración, mostrando serenidad y reconduciendo la actividad.
- *Motivación.* Se evalúa el entusiasmo que muestra el alumno ante las tareas y los materiales.
- *Flexibilidad.* Se evalúa la capacidad de generar alternativas de solución a las tareas.
- *Interactividad.* Se evalúa la capacidad de mantener intercambios comunicativos con cierto grado de elaboración durante el desarrollo de las tareas.
- *Responsividad.* Se evalúa la actitud del alumno ante el aprendizaje y ante las propuestas del evaluador.

La aplicación completa del ACFS puede durar unas dos horas. Los autores recomiendan que su administración se realice en dos sesiones, incluyendo tres subescalas cada vez, e incluso, en función de las características de los niños, puede ser realizada una subescala por sesión. Por otra parte, un aspecto muy interesante de esta prueba con respecto a la evaluación de alumnos que presentan dificultades en el lenguaje oral es que, aunque algunas de las tareas, así como la valoración que se efectúa en función de los indicadores cualitativos de conducta, tienen una importante base verbal, la prueba puede ser utilizada con niños que presentan problemas o retraso en el lenguaje, ya que, por una parte, otras tareas que conforman la prueba pueden ser desarrolladas sin la utilización del lenguaje oral y, por otra, su aplicación puede ser muy valiosa para valorar la calidad de la respuesta a la intervención en estos niños y la evolución de su desarrollo.

La prueba ACFS no es una prueba normalizada. Las medidas obtenidas son cuantitativas y también cualitativas. Las puntuaciones cuantitativas, compuestas por el número de ítems correctos y los porcentajes de aciertos, indicarían el nivel de competencia alcanzado en cada una de las tareas propuestas, pero no deberían servir según sus autores, como indicadores de equivalencia en función de la edad. La interpretación cuantitativa tiene una función de calibrar la evolución del desarrollo de los alumnos evaluados, además de su utilidad a efectos de investigación; sin embargo, la mayor validez de esta prueba para la evaluación parece residir en que puede otorgar información cualitativa acerca de la puesta en práctica de determinadas funciones cognitivas durante el desarrollo de tareas de aprendizaje, así como sobre la capacidad de interactuar y aprender en una situación colaborativa.

Investigación y revisión crítica del modelo

Lo primero que debemos decir sobre la investigación realizada desde esta perspectiva, es que los estudios desarrollados son todavía pocos y, así mismo, los datos que se pueden extraer de estos primeros estudios sólo apuntan algunas conclusiones que, en cualquier caso, no parecen todavía sólidas. Estos primeros estudios se han realizado básicamente para establecer las primeras propiedades psicométricas de la escala. En este sentido, un primer aspecto relevante de la valoración de la efectividad de un procedimiento de evaluación dinámica que siga un formato pretest/ mediación /postest, es la comprobación de que existen diferencias (ganancias) entre la aplicación del pretest

y la del postest. Esta comprobación informaría de entrada sobre la validez de constructo de la prueba. En un primer estudio realizado a modo de aplicación piloto (Lidz, 2000), treinta niños de una escuela infantil privada de N. York y con un previsible nivel alto de funcionamiento cognitivo, fueron evaluados con la ACFS. Este estudio, a parte de otorgar información sobre las diferencias en las puntuaciones entre las fases pre y postest, sirvió para retocar algunas de las subescalas de la prueba en función de algunos de los resultados obtenidos; de esta forma, de las cinco originales se pasó a las seis subescalas que componen hoy la prueba, ya que la tarea de *Planificación* fue rediseñada y descompuesta en dos: *Planificación verbal* y *Perspectiva*. También fue rediseñada la tarea de *Realización secuencial de patrones*. Se obtuvieron ganancias significativas para tres de las subescalas: *Clasificación*, *Memoria auditiva inmediata* y *Memoria visual inmediata*. Uno de los aspectos interesantes de este estudio es que el contenido de la prueba ACFS fue revisado en función del contexto de aplicación del estudio. De esta forma, al tratarse de alumnado con un alto nivel de funcionamiento cognitivo, algunas de las tareas de la escala debieron ser adaptadas «al alza» para evitar el efecto techo.

En otro estudio llevado a cabo por Brooks (1997) en el Reino Unido, se usó sólo el subtest de *Clasificación* de la prueba ACFS con una muestra de 22 niños de educación infantil que presentaban trastornos del desarrollo. La mitad de los niños fue asignada al GE y la otra mitad al GC. El objetivo del estudio se centraba en determinar el posible efecto de práctica de la prueba; de este modo, el GC recibió el pretest y el postest, pero no recibió la intervención, mientras que el GE recibió todo el proceso. Tras la evaluación inicial, la totalidad de los niños de la muestra había sido descrita como constructores o «apiladores», dado el comportamiento que mantenían en interacción con los bloques proporcionados en la prueba. Los resultados mostraron que dos tercios de los niños del GE pasaron de ser considerados «apiladores» a «agrupadores» en función de su conducta con los bloques, mientras que ninguno del GC pasó a engrosar la nueva categoría.

En otro estudio llevado a cabo en N. York con 26 niños de 4 años, de los cuales todos menos cinco presentaban trastornos del desarrollo, Shurin (1998) encontró ganancias significativas en cuatro de las seis subescalas de la prueba: *Clasificación*, *Perspectiva*, *Planificación* y *Realización secuencial de patrones* y, al contrario de lo que ocurriera en el primero de los estudios presentados, no halló diferencias en los dos subtests de *Memoria*. La propia Lidz (2000) apunta a la posibilidad de que déficits de memoria asociados a los trastornos del desarrollo pudieran estar en la base de los resultados. En este estudio, Shurin evaluó los indicadores cualitativos de conducta, a raíz de los cuales se efectuaron distintos análisis; por una parte, se trató de determinar la correlación que mantenía la puntuación total de la ACFS con la puntuación que era posible obtener de los indicadores. La correlación fue muy significativa ($r = 0,65$; $p < 0,001$), por lo que se podía deducir que la actuación en la escala estaba relacionada con la puesta en práctica de los indicadores valorados. Posteriormente se quiso establecer qué indicadores correlacionaban con la puntuación obtenida en las tareas de la escala. Las correlaciones fueron significativas en este caso para la *Persistencia* y la *Flexibilidad*, estando al borde de serlo también para la *Tolerancia a la frustración*. Por otra parte, el análisis de la consistencia de la escala de indicadores reveló que todos ellos estaban correlacionados de manera aceptable con la puntuación total menos la *Interactividad*, cuya correlación, sin ser significativa estaba al borde de serlo ($r = 0,37$; $p < 0,06$).

El estudio de Levy (1999) sobre la validez discriminante de la ACFS se realizó mediante la aplicación de la prueba a un grupo de 22 niños con edades comprendidas entre los 4 años y los 4 años y 11 meses. La mitad de la muestra estaba diagnosticada con moderados retrasos en su desarrollo, especialmente en tareas de Lenguaje y cognición en general. Los resultados en la evaluación pretest arrojaron diferencias a favor de los niños sin dificultades en todas las subescalas, las cuales eran significativas solo en dos de ellas y en la puntuación total; sin embargo, tras la fase de intervención, los resultados en el posttest mostraron diferencias significativas a favor del mismo grupo en cuatro de los subtest además de la puntuación total. El grupo con dificultades en su desarrollo había experimentado ganancias solo en una subescala (Memoria visual) y en la puntuación total, mientras que el otro grupo presentó ganancias significativas en tres de las subescalas y también en la puntuación total. El estudio apunta a que pudo producirse en este grupo sin dificultades un cierto efecto techo en algunas de las subescalas, dado que las puntuaciones en el pretest fueron muy altas. En definitiva, como puede observarse, lo que confirma igualmente el propio análisis de Lidz sobre este estudio, el análisis de la validez de la prueba para discriminar en función de la agrupación previa efectuada por los Servicios de Orientación de la escuela, reveló una potencia moderada. En nuestra opinión, aunque la información de la que disponemos es limitada, es posible, como así se recoge en la literatura sobre la aplicación dinámica, que el proceso de evaluación seguido difumina diferencias entre ambos grupos que, de otra forma y con otro tipo de evaluación más convencional, aparecerían probablemente con más claridad (Budoff, 1968; Feuerstein, 1991; 1998a; Sternberg y Grigorenko, 2003; Resing, 2000). Esto, a nuestro parecer, lejos de ser un resultado negativo apoyaría la tesis de que la intervención ha logrado acercar o, al menos, disminuir las diferencias entre grupos que partían de posiciones muy distintas a priori. En cualquier caso, el objetivo esta vez era otro para los investigadores, que se proponían analizar la validez de discriminación de la prueba; en este sentido, nos parece que los problemas metodológicos derivados de lo exiguo de la muestra y del identificado efecto techo en el grupo sin dificultades, hace difícil extraer una conclusiones del todo válidas sobre la potencia discriminadora de la prueba.

Por último, algunos estudios también han abordado la fiabilidad de las observaciones y las puntuaciones sobre los indicadores cualitativos de conducta en la ACFS. Por ejemplo, en el estudio citado anteriormente de Shurin (1998) se informó de un acuerdo inicial inter-observadores del 74% para el pretest y del 82% para la fase de mediación. Posteriormente, una vez comparadas y combinadas las valoraciones, el acuerdo fue del 96%. En otro estudio de los citados (Levy, 1999), se alcanzó un porcentaje de acuerdo total bastante menor (67%) una vez combinadas y valoradas conjuntamente las observaciones del pretest y de la intervención. En un nuevo estudio (Aranov, 1999) se evaluó la fiabilidad inter-observadores para la valoración de los indicadores con una muestra de 25 niños de cuatro clases de educación infantil que presentaban necesidades educativas especiales. Las valoraciones del investigador fueron comparadas con los profesores del aula y con los logopedas del centro. Los resultados mostraron correlaciones bastante altas en todas las subescalas (entre 0,58 y 1,00) con los profesores, siendo la correlación de 0,95 para la puntuación total; y aún más altas con los logopedas (entre 0,78 y 1,00) en las subescalas y 0,97 para el total. El estudio también indicó un coeficiente Alfa de fiabilidad de 0,77 para la escala de indicadores observada por el investigador y de 0,81 para la observada por uno de los logopedas.

Como hemos podido constatar, la prueba desarrollada por Lidz ha realizado aún pocos estudios que puedan avalar sólidamente aspectos psicométricos de validez y fiabilidad. En este sentido, los estudios presentados han sido realizados con muestras en algunos casos excesivamente pequeñas y, en algunos casos, sin contar con grupos de control adecuados. Por otra parte, no conocemos ningún estudio que haya abordado lo que quizás constituye el aspecto más importante de un instrumento como éste, es decir, en qué medida, una prueba dinámica basada en el currículum como ACFS es capaz de obtener una mayor validez predictiva sobre el rendimiento escolar que otras pruebas dinámicas ajenas a los programas escolares. Estos estudios necesitarán sin duda de diseños más amplios que cuenten con el desarrollo del currículum escolar y que valoren posteriormente, tanto las predicciones realizadas sobre el desempeño de los niños, como la incidencia de las orientaciones psicopedagógicas otorgadas al profesorado a raíz de la valoración dinámica. Sin duda, los resultados de estos estudios serán de nuestro interés.

La prueba ACFS presenta a nuestro entender algunas limitaciones de las que aún no hemos hecho mención. Una de ellas podría relacionarse con la ausencia de valoración del contexto de enseñanza y aprendizaje; la otra, más relacionada directamente con los contenidos de la prueba, hace referencia a la ausencia igualmente de valoración del lenguaje escrito. En este sentido, quizás la inclusión de tareas relacionadas con la conciencia fonológica, sobre todo a partir de los cuatro años, podría ofrecer información muy valiosa a efectos de establecer posibles dificultades relacionadas con los procesos metalingüísticos, que a su vez podrían predecir futuros problemas en el acceso adecuado a los procesos de decodificación y comprensión. En cualquier caso, a pesar de la relevancia que suponen estos dos aspectos para establecer una evaluación contextualizada y que recoja contenidos significativos en el desarrollo educativo del alumnado de educación infantil, sobre todo del alumnado de 4 y 5 años, la ACFS parece caminar en una dirección de alguna forma cercana a la que más tarde propondremos nosotros. Contextualizar los instrumentos de evaluación, analizando los contenidos que forman parte del currículum escolar y actuando sobre las funciones cognitivas, metacognitivas y motivacionales que subyacen al aprendizaje y desarrollo de esos contenidos, nos parece a nosotros también una forma idónea de aproximarnos a la evaluación.

La evaluación en la Zona de Desarrollo Potencial y el método de las sugerencias graduadas de J. C. Campione y A. Brown

Bases teóricas y metodológicas

Campione, Brown y Ferrara (1987) concebían la inteligencia en términos de eficiencia en la capacidad para interiorizar nuevos aprendizajes y para aprovecharse de formas de instrucción incompleta. A su vez, estos autores consideraban la inteligencia en estrecha relación con la capacidad de transferir el aprendizaje a nuevas situaciones. A través de esta conceptualización, que servía de marco a sus propuestas educativas, así como al método de aplicación dinámica desarrollado por ellos, la capacidad de aprender se relacionaba estrechamente con la metacognición y con la capacidad de aplicar los conocimientos y procesos aprendidos a nuevos contextos de aprendizaje (Resing, 2000). Otros aspectos considerados relevantes de la inteligencia, como la capacidad de adaptación eficaz al medio, que han sido especialmente destacados por otros autores, entre ellos Feuerstein, no se recogían en la definición planteada; si bien es cierto que podría entenderse que estos conceptos relacionados con la adaptación continua y

flexible al entorno, se encuentran englobados en su idea de la capacidad de interiorizar nuevos aprendizajes (Resing, 1998).

Lógicamente, esta consideración dinámica de la inteligencia, como capacidad de aprendizaje, es la justificación para el empleo de pruebas o actividades de evaluación dinámica en la valoración de este constructo. Por otra parte, otro aspecto que se desprende igualmente de la concepción de la inteligencia propuesta por estos autores, es el contenido de la prueba dinámica que conforma su método de las sugerencias graduadas; en éste son parte fundamental la evaluación de la capacidad del sujeto para interiorizar y aplicar estrategias sugeridas por el evaluador en tareas de razonamiento inductivo o aritmética, así como la evaluación de la transferencia a otras tareas presentadas.

Este procedimiento de evaluación dinámica fue desarrollado en total sintonía con el marco teórico de los estudios de Vygotski sobre la ZDP. Los autores establecen una situación de prueba dinámica en la que el sujeto recibe un apoyo gradual en la resolución de una actividad hasta que se muestra capaz de realizarla por sí mismo (Campione y Brown, 1987; Campione, 1989). La cantidad de sugerencias necesarias para dominar una tarea se convierte en una de las medidas que sirven para operativizar el método; la otra medida utilizada tiene que ver con la transferencia, es decir, con la capacidad que muestra el sujeto para poner en práctica los aprendizajes adquiridos en situaciones de aprendizaje distintas a aquellas en las que fueron aprendidos. Esta capacidad de transferir o transcontextualizar (Pozo, Monereo y Castelló, 2001) el aprendizaje se ha revelado clave en los procesos educativos e incluso ha sido utilizada como elemento discriminante a raíz del cual puede decirse que ha tenido lugar un verdadero aprendizaje (Campione y Brown, 1977; Pozo, 1996). La transferencia es cuantificada por los autores mediante la puntuación en la realización de tareas similares a las tareas de instrucción, pero que van diferenciándose progresivamente de éstas. La operativización, por tanto, se realiza con base en el número de sugerencias necesarias para que el sujeto realice por sí mismo la tarea y en el número de sugerencias que requiere para resolver por su cuenta los problemas de transferencia planteados.

Sin duda este enfoque introduce una novedad importante en el desarrollo de la evaluación dinámica, ya que, por una parte, focaliza la atención en la tarea y en la cantidad y tipo de ayuda necesaria para resolverla, más que en el rendimiento exclusivo del sujeto. Por otra parte, introduce un análisis cuantificado y normalizado de la prueba. La medida resultante se considera en relación con la cantidad de ayuda necesaria para alcanzar un aprendizaje y para transferirlo a otras situaciones. Esta medida es interpretada por los autores como un potencial del aprendizaje (Sternberg y Grigorenko, 2003), es decir, un indicador de la ZDP del alumno. En este sentido, los niños y niñas que necesitan menos ayuda para resolver las tareas propuestas, se caracterizarían por una mayor amplitud de su ZDP.

Decíamos en páginas anteriores que las pruebas de evaluación dinámica podían encuadrarse básicamente en dos grandes paradigmas en función de su diseño metodológico y, más concretamente, en función de las fases de las que constan en el desarrollo de la prueba. Uno de estos modelos lo constituían las pruebas que respondían, en mayor o menor medida, a un modelo pretest/ mediación /postest y el otro, aunque más difusamente, respondía a un modelo pretest/ tarea  retroalimentación continua /postest. Aunque ambos paradigmas comparten en su diseño la presencia de una

evaluación previa y otra posterior a la intervención realizada, la característica esencial del método de las sugerencias graduadas, por lo cual encarnaría este segundo modelo, podríamos situarla en la fase de intervención, que constituye en realidad el momento relevante del proceso y es donde se obtiene y operativiza la información que determinará la medida resultante. La adecuación del modelo metodológico desarrollado por Campione y Brown al esquema teórico propuesto por Vygotski sobre la ZDP y sobre el modo en que podría tener lugar una evaluación de aquellos procesos que están en desarrollo en el sujeto, accediendo de esta forma a una valoración del nivel potencial del aprendizaje, se trata de concretar en la contemplación de una secuencia que consta de cuatro momentos diferenciados: (1) recogida de información sobre el nivel de actuación inicial del sujeto sin ayuda (pretest), (2) fase de aprendizaje inicial, en la que tienen lugar las sugerencias dadas por el evaluador, tanto en las tareas de aprendizaje como en las de transferencia, (3) comprobación estática del mantenimiento del aprendizaje y la transferencia (postest), y (4) comprobación dinámica del mantenimiento del aprendizaje y de la transferencia (postest con mediación). Se considera que el procedimiento es efectivo si la actuación del sujeto mejora con la instrucción, si el efecto de esta instrucción es duradero y si el resultado del aprendizaje puede ser transferido a tareas distintas de aquellas en las que fue aprendido (Brown y Campione, 1981; Brown, Bransford, Ferrara y Campione, 1983).

En cuanto a la población para la que este método se ha desarrollado, ésta está compuesta principalmente por alumnado que presenta bajo rendimiento escolar, alumnado con DA o deficiencia mental leve y moderada (Campione y Brown, 1987). Las tareas empleadas incluyen problemas de razonamiento inductivo, como variaciones de las matrices progresivas de Raven y tareas de completamiento de series; también incluyen problemas matemáticos y tareas de comprensión auditiva y de lectura.

Investigación y revisión crítica del modelo

La mediación efectuada por los aplicadores se realiza mediante sugerencias previamente determinadas, que son dadas al sujeto cuando éste las requiere para poder resolver una tarea. La estructuración del proceso de mediación es aquí mayor que en el enfoque original de Feuerstein; la tarea y las sugerencias desarrolladas prevén de antemano la práctica totalidad de opciones que, en su búsqueda de soluciones, puede adoptar el sujeto. Esta mayor estructuración tiene ciertas ventajas asociadas a la posibilidad de cuantificar precisamente el proceso más relevante en una prueba de evaluación dinámica, es decir, la interacción mediada, pero plantea también, obviamente, algunos inconvenientes. Entre éstos, Lidz (1991) destaca que la distinta naturaleza de la variable de medida, esto es, las sugerencias aportadas, puede poner en cuestión el hecho de que los resultados obtenidos con base en distintos tipos de sugerencias sean comparables y aditivos (Saldaña, 2001). Además podría entenderse que las sugerencias pueden tener también distinto efecto y significado en función del perfil cognitivo del sujeto, confundiéndose por tanto la medida del potencial de aprendizaje con características como la capacidad de atención y reflexión en las tareas (Sternberg y Grigorenko, 2003). Este aspecto, que por otra parte podría ser aplicado a la gran mayoría de las pruebas diseñadas, puede ser realmente difícil de solucionar, ya que no parece posible separar hasta tal punto los procesos y el perfil cognitivo que tienen lugar en el desarrollo de actividades de mediación; por otra parte, los efectos de estas variables podrían atenuarse contando con muestras compuestas por sujetos «iguales» en función de características

tales como: dificultades de aprendizaje, bajo rendimiento cognitivo, estilos de aprendizaje, etc.

Otro problema con el que se encuentra este procedimiento cuando se ha acometido el análisis de su validez, tiene que ver más directamente con la cuantificación de los resultados y con el modo en que se lleva a cabo. De este modo, Sternberg y Grigorenko (2003) plantean en su revisión que el método de las sugerencias graduadas podría llegar a ser interpretado como una sucesión de pequeñas tareas en las que el número de sugerencias ofrecidas puede ser visto como el número de subtareas que el niño no es capaz de resolver por su cuenta; así pues, la prueba podría puntuarse con base en las respuestas correctas del alumno en lugar de hacerlo con base en las sugerencias. Según estos autores, la prueba, vista de esta forma, podría ser interpretada desde una concepción más estática que dinámica. Esta última consideración merece una reflexión por nuestra parte, por cuánto nos parece relevante e interesante el análisis efectuado y, así mismo, por cuánto nos parece hasta cierto punto contradictorio el énfasis que estos mismos autores hacen sobre la necesidad de establecer medidas cuantificables y normalizadas en las pruebas de evaluación dinámica y la aparente imposibilidad que esto plantea, ya que haciendo un cierto ejercicio de generalización y extrapolación del análisis realizado por ellos a cualquier prueba de evaluación, en la medida en que ésta asuma un paradigma cuantitativo para la valoración de la información obtenida, podría ser interpretada como una sucesión de subtareas y, por tanto, más fácilmente encuadrada en modelos estáticos. En definitiva, esto plantea un grave problema para la evaluación dinámica que, por una parte, busca y hasta cierto punto necesita diseños y modelos teóricos y metodológicos que sean válidos y fiables en un uso generalizado de los mismos, y, por otra, constata que los procesos de cuantificación y normalización de las pruebas conllevan casi irremediabilmente pérdidas o renunciadas a la consideración plena de elementos significativos de su propia conceptualización.

Los resultados de algunas de las investigaciones efectuadas en torno a la aplicación del método de sugerencias graduadas ofrecen algunas conclusiones que nos parecen especialmente relevantes. En primer lugar, el estudio de los procesos de aprendizaje y la transferencia en alumnado con distinto nivel de rendimiento académico y también con distinto nivel educativo (curso) y edad, concluyen que el número de sugerencias necesarias para resolver las tareas de aprendizaje y, sobre todo para resolver las tareas que requieren una mayor transferencia, aumentaba significativamente en alumnos con un menor rendimiento académico y en alumnos de menor edad (Campioni et. al., 1985; Ferrara, Brown y Campione, 1986). En esta misma línea, otros estudios han destacado y cuantificado las diferencias en la actuación de niños con y sin déficit cognitivo, medidas a través del número de sugerencias necesitadas (Campioni y Brown, 1987; Sternberg y Grigorenko, 2003). Más allá de lo esperado de tales resultados, un aspecto muy relevante de estos estudios, es que apunta a la consideración de diferencias evolutivas en los procesos de aprendizaje y transferencia. Estas diferencias pueden ser interpretadas como un mayor nivel de procesamiento de la información alcanzado por los alumnos mayores (Klar, 1992; Karmiloff-Smith, 1994), por lo que necesitarían menos sugerencias para poder dar solución a las actividades planteadas, siendo congruente esta interpretación de los resultados con la que se propone como posible en otras investigaciones relacionadas con el conocimiento metacognitivo (Peronard, Crespo y Velásquez, 2000; Peronard, Crespo y Guerrero, 2001), entre ellas la llevada a cabo por nosotros mismos acerca de los metaconocimientos sobre la comprensión de textos en muestras de alumnado con y sin DA (Navarro y Mora, 2003a), en la que se perfila un

cierto patrón evolutivo en la adquisición de los metaconocimientos relacionados con la lectura. Otra interpretación plausible de los resultados obtenidos en la investigación con este método es que el potencial de aprendizaje, tal como también destacábamos cuando analizábamos el modelo de Feuerstein, no sería un elemento estático en la configuración cognitiva de una persona; por el contrario, se trataría de un elemento que se desarrolla de manera dinámica. Dicho de otro modo, la «ampliación», o el traslado continuo hacia delante, de la ZDP tendría lugar en función de las experiencias de aprendizaje mediado —utilizando la terminología de Feuerstein— que, a su vez están habitualmente relacionadas con el desarrollo evolutivo de los sujetos.

En otros estudios realizados por estos investigadores se ha puesto de manifiesto la relevancia del proceso de transferencia y del mantenimiento del aprendizaje. En un estudio comparativo realizado con niños que partían con una «edad mental» equivalente, pero que diferían en su consideración de discapacidad intelectual, Campione, Brown, Ferrara, Jones y Steinberg (1985) utilizaron la prueba basada en el test de matrices progresivas de Raven para determinar posibles diferencias en sus niveles de actuación. Los resultados mostraron que, aunque en la fase de aprendizaje inicial no había diferencias significativas entre ambos grupos, éstas comenzaban a aparecer cuando se pasaba a la fase de evaluación del mantenimiento del aprendizaje y de la transferencia, en las que las puntuaciones obtenidas por los alumnos con discapacidad intelectual fueron menores. También mostraron que a medida que las tareas que evaluaban la transferencia se alejaban de las iniciales y requerían por tanto una mayor flexibilidad en la aplicación de las estrategias, los alumnos con discapacidad puntuaban significativamente por debajo de sus compañeros.

En cuanto a la investigación relacionada con los procesos de lectura, Day y Zajakowski (1991) pusieron en marcha un estudio en el que compararon la actuación con y sin mediación de niños que presentaban una aptitud media para la lectura, con niños que presentaban dificultades en su lectura. Encontraron que los niños con problemas de lectura necesitaban un número significativamente superior de sugerencias para adquirir el dominio en la tarea criterio de lectura.

Otro aspecto destacable de este enfoque es la relevancia del componente de instrucción metacognitiva en el proceso de evaluación dinámica. En efecto, Resing (1993), a cuyo trabajo dedicaremos un apartado especial posteriormente, encontró importantes diferencias significativas entre un grupo de alumnos que había recibido instrucción metacognitiva en el marco del método de sugerencias graduadas, y el grupo de control; también informó sobre diferencias significativas en cuanto al grado en que predecían el rendimiento escolar las puntuaciones del potencial de aprendizaje (derivadas del proceso de aprendizaje y transferencia) y las puntuaciones obtenidas en el postest, con respecto a las puntuaciones del pretest.

En relación a la capacidad predictiva del método y la información ofrecida a los evaluadores, los autores afirman que las puntuaciones de aprendizaje y transferencia ofrecen una información más rica que las pruebas de carácter estático (Campione, 1989), por cuanto las puntuaciones de mejora entre el pretest y el postest eran mejor predichas por la medida del potencial de aprendizaje obtenido a raíz de las puntuaciones de aprendizaje guiado y de transferencia en la fase de intervención, que por las puntuaciones correspondientes al nivel de desarrollo inicial del sujeto. Se han realizado varios estudios para evaluar la contribución de las puntuaciones de aprendizaje y

transferencia, así como de otros factores como el CI y las puntuaciones pretest, en la mejora observada entre pretest y postest. Las conclusiones de algunos de estos estudios, aunque no parecen concluyentes, apuntan a que, tal como señalábamos anteriormente, la puntuación del potencial de aprendizaje predecía mejor la mejora observada que las pruebas estáticas de aptitud y que las puntuaciones pretest (Campione y Brown, 1987; Campione, 1989).

Aunque la solidez de los resultados obtenidos en estas y otras investigaciones llevadas a cabo por los autores y por otros investigadores interesados en el método de sugerencias graduadas, no parece aún suficiente (Sternberg y Grigorenko, 2003) y deben abordarse de forma más concluyente aspectos nada desdeñables relativos, por ejemplo, a la coherencia interna del método, la fiabilidad test-retest, así como a la validez predictiva de sus medidas del potencial de aprendizaje con relación a criterios externos como el rendimiento escolar; el modelo ofrecido para concretar la evaluación de y en la ZDP, basado en la retroalimentación continua que es ofrecida al sujeto a través de pistas o sugerencias que apoyan la realización de una tarea, ha supuesto quizás el primer intento real de análisis cuantitativo del potencial de aprendizaje. Los datos aportados por la investigación, aunque limitados, han abierto nuevas vías de investigación y aplicación dinámica, y han contribuido decisivamente a potenciar las razones y argumentos que se esgrimen a favor de la utilización de la evaluación dinámica en la educación escolar.

La evaluación del potencial de aprendizaje a través del razonamiento inductivo (LIR) de Wilma C.M. Resing

Bases teóricas y metodológicas

Quizás ha sido Resing quien ha seguido más de cerca la estela dejada por Campione y Brown al plantear la evaluación del potencial de aprendizaje desde una perspectiva más centrada en la tarea y en la ayuda necesaria para que ésta pueda llegar a ser resuelta con éxito. Esta autora, partiendo de una concepción de los procesos cognitivos y la inteligencia deudora de los trabajos de Vygotski (1934/1995) y de los de Brown y Campione (Campione y Brown, 1987; Brown y Ferrara, 1985; Campione, Brown y Ferrara, 1987; Campione, Brown, Ferrara y Bryant, 1984; Campione, Brown, Ferrara, Jones y Steinberg, 1985; Ferrara, Brown y Campione, 1986) y partiendo de la idea básica del traspaso del control y la regulación de la actividad mental desde una regulación interpsicológica a una intrapsicológica, se propone evaluar la capacidad de razonamiento inductivo en niños pequeños (7-8 años) y valorar cómo estos niños son capaces de aprender este tipo de razonamiento a través de la exposición a un periodo relativamente corto de instrucción. Según la autora, el hecho de que su propuesta de evaluación se dirija a niños de la edad citada, se relaciona con la relevancia que suelen tener en esos momentos determinadas decisiones relacionadas con la modalidad de escolarización apropiada y, más concretamente, con relación a las decisiones sobre la necesidad de una educación especial. Por otra parte, el uso de actividades de razonamiento inductivo ha sido muy extendido entre los investigadores de procesos de evaluación dinámica. Estas tareas, como las matrices progresivas o las series numéricas que deben ser completadas, parecen prestarse muy bien a la mediación que intenta provocar una actividad constructiva y de más alto nivel cognitivo en los sujetos (Resing, 2000; Sternberg y Grigorenko; 2003) La prueba desarrollada por Resing (LIR) se basa principalmente en dos tipos de actividades para evaluar el razonamiento inductivo: *exclusión visual y analogías verbales*.

El procedimiento desarrollado por Resing en su prueba difiere del propuesto por Campione y Brown en que el formato de evaluación que se utiliza es el de pretest/mediación/postest; eso sí, en este caso, la mediación está compuesta por sugerencias graduadas hasta que el alumno es capaz de resolver por sí mismo la actividad. Así mismo, al igual que en el método de Campione y Brown, la cantidad y el tipo de estas sugerencias determinarán el potencial de aprendizaje del sujeto. De esta forma, el grado de ayuda necesaria para alcanzar el criterio de aprendizaje determinado previamente, puede ser entendido como un indicador de la estrechez o amplitud de la ZDP del niño (Resing, 2000). Por otra parte, la prueba LIR podría ser considerada una revisión «dinámica» de una anterior prueba desarrollada por Bleichrodt, Drenth, Aaal y la propia Resing (1984) para la evaluación de la inteligencia en los niños (*Revised Amsterdam Child Intelligence Test -RAKIT*). En esta prueba, los niños debían resolver de forma independiente dos tipos de actividades de razonamiento inductivo; una no verbal, en la que debían descubrir la regla que subyacía a la relación existente entre tres figuras geométricas abstractas y que excluía una cuarta figura, la cual debía ser señalada y seleccionada por el niño; y otra que consistía en una prueba de analogías verbales, en la que los niños debían inferir la relación existente entre los dos primeros términos de la analogía para aplicar esa relación a un segundo par de términos del que sólo se disponía del primero de ellos, debiendo seleccionar el segundo de entre cuatro posibles opciones. Se incluían dibujos de cada uno de los términos que aparecían en las analogías, ya que el test podía ser aplicado a niños con problemas en lectura o en el recuerdo de palabras. De entre las analogías verbales más frecuentemente usadas en los test, y especialmente en el RAKIT (analogías funcionales, parte-todo, analogías de conversión, de pertenencia a una clase, de localización, etc.), Resing decidió construir su prueba dinámica centrándose en dos tipos: analogías con una relación funcional y analogías con una relación parte-todo. En el cuadro 14 podemos ver algunos ejemplos de los tipos de analogías usados en el LIR.

Cuadro 14. Tipos y ejemplos de relaciones de analogías verbales usadas en las sesiones de instrucción de la evaluación del potencial de aprendizaje a través del razonamiento inductivo (LIR) Tomado de Resing (2000).

Tipo de relación	Definición	Ejemplos
<i>Relación funcional</i>	Un término de la analogía cumple o realiza una función para el otro término	oído: oír = ojo: ver zapato: pie = sombrero: cabeza
<i>Relación parte-todo</i>	Un término de la analogía es parte integrante del otro término	elefante: trompa = hombre: nariz camisa: botones = zapatos: cordones

De la misma forma, los ítems de razonamiento inductivo de exclusión o identificación visual también han sido construidos por Resing con base principalmente en el test RAKIT. En los nuevos ítems, los principios seguidos para la transformación de las figuras que han de ser relacionadas, han sido los de transformación de la posición y transformación de la forma. Para cada uno de estos tipos de transformación, las figuras pueden sufrir rotaciones, cambios de tamaño, de forma o la combinación de estos distintos tipos de cambio. En el siguiente cuadro recogemos resumidamente los tipos y relaciones de transformación que seguían las figuras de los ítems presentados.

Cuadro 15. Tipos y definiciones de relaciones de exclusión visual usadas en las sesiones de instrucción del LIR. Tomado de Resing (2000).

Principios de construcción de los ítems	Definiciones
Principio de exclusión de una figura	Principios de transformación comunes a cada una de las cuatro figuras
Cambio de posición	Rotación - Tamaño - Forma - Rotación y Tamaño Rotación y Forma - Tamaño y Forma
Cambio de forma	Rotación - Tamaño - Forma - Rotación y Tamaño Rotación y Posición - Tamaño y Posición

Por otra parte, como decíamos en el anterior apartado, cuando hacíamos referencia a uno de los estudios de la propia Resing (1993), el componente metacognitivo en la fase de instrucción es considerado esencial. Las estrategias de control ejecutivo como planificar, realizar comprobaciones, monitorizar y supervisar las propias realizaciones o evaluarse, son consideradas clave para que se pueda producir un efecto duradero del aprendizaje obtenido y éste pueda ser transferido y generalizado a situaciones de aprendizaje distintas (Campione, Brown y Ferrara, 1987). El proceso seguido durante esta fase de instrucción incluye varios pasos en los que el sujeto que lo requiere recibe una serie de sugerencias que buscan la aplicación de estrategias cognitivas y metacognitivas para poder solucionar el problema. En el siguiente cuadro, tomado de Resing (2000), podemos ver esta secuencia de forma más clara.

Cuadro 16. Representación esquemática de la estructura de sugerencias graduadas para la fase de instrucción de analogías verbales «parte-todo», con una estructura $A:B = C: ?$ (D1-D5)

Sugerencias metacognitivas	Sugerencias cognitivas	Graduación en la fase de instrucción
Acotar el problema		¿Qué es lo que tienes que hacer?
Reconocer el problema		¿Habías hecho esto antes?
Activar el conocimiento	Codificar	Piensa en las dos primeras palabras (Ay B)
Selección de la estrategia	Inferencia	¿Por qué crees que estas palabras van juntas? (A y B)
	Codificar	Piensa en la primera palabra de la segunda pareja (C)
Planificar	Aplicación del plan trazado	Las dos primeras están juntas porque A es parte de B; entonces, ¿qué palabra seguirá a C?
Monitorizar/Supervisar	Comparación	¿Va D1 (la primera de las 4 opciones) con C? (etc.)
Comprobar/Evaluar	Justificación de la R.	
Comprobar sistemáticamente la R.	Dar la respuesta	
Aprendizaje de la retroalimentación		Evaluar las respuestas dadas

Cada sujeto recibe seis sesiones de instrucción o entrenamiento que siguen a la realización del pretest. Cada sesión tiene lugar hasta que el sujeto consigue responder adecuadamente a las tareas propuestas sin la ayuda del evaluador. Para que una respuesta sea considerada como correcta, no sólo se debe elegir la alternativa adecuada de entre las opciones disponibles, sino que se debe justificar la elección, mostrando de este modo hasta qué grado de comprensión de la actividad ha llegado el sujeto.

Investigación y revisión crítica del modelo

Los estudios realizados por Resing y sus colaboradores para evaluar su propuesta fueron realizados con un total de 325 alumnos y alumnas, 175 fueron asignados al GE y 150 al GC. A los sujetos que integraban el GE les fue aplicado el instrumento en todas sus fases e igualmente, a los de GC les fue administrado la totalidad del LIR, incluyendo los contenidos de la fase de entrenamiento, aunque, en este caso, fueron entrenados mediante un método de enseñanza tradicional o convencional, en lugar de seguir la secuencia de las sugerencias graduadas. La muestra de estudio estuvo compuesta por alumnado de educación especial, alumnado con dificultades de aprendizaje, de minorías étnicas y alumnado que no presentaba dificultades en su aprendizaje (Resing, 1990; 1993; 1997; Resing y Van Wijk, 1996; Resing, Slenders, Bosma y Iedema, 1997; Slenders y Resing, 1997). Para valorar la aportación del LIR a la predicción del rendimiento escolar con relación a una prueba estática, se comparó la validez predictiva que otorgaban las puntuaciones en la fase de entrenamiento y las del postest en la prueba de Resing, con la validez en la predicción obtenida a raíz de las puntuaciones obtenidas con el test de inteligencia RAKIT. Los resultados mostraron que, tanto las puntuaciones postest del LIR como las puntuaciones de la fase de entrenamiento, aportaban una validez predictiva adicional de 0,15 sobre la base de correlación de 0,60 que mantenía el RAKIT con el rendimiento escolar, medido a través de las calificaciones escolares y la valoración del profesorado. La conclusión que se hacía de estos análisis era que la evaluación dinámica llevada a cabo mediante la prueba LIR realizaba una contribución adicional a la información obtenida con un test de inteligencia empleado convencionalmente.

El análisis cualitativo, como extensión del realizado en función de aspectos psicométricos como la fiabilidad, la validez o el valor de predicción de la prueba, permite extraer información, a nuestro juicio, especialmente valiosa. Por ejemplo, se encontró que la cantidad de ayuda necesitada para resolver las tareas con éxito, así como las diferencias entre las puntuaciones obtenidas en el pretest y el postest, establecían importantes diferencias entre los alumnos de la muestra; así mismo, permitió analizar casos individuales; así, en la fase de entrenamiento se comprobó que algunos alumnos necesitaban mucha más ayuda para resolver uno de los dos tipos de tareas, analogías o exclusión visual, a pesar de que la correlación entre estas tareas era muy alta: $r = 0,73$. Analizadas las ayudas necesitadas sobre la base de las puntuaciones en el pretest, es decir, sin intervención, el número de ayudas que los alumnos necesitaban no se correspondía con lo esperado en función de estas puntuaciones; de esta forma, alumnos con bajas puntuaciones estáticas necesitaban menos ayuda de lo esperado y algunos alumnos con altas puntuaciones necesitaban mayor número de sugerencias de lo que se esperaba. De este análisis podría derivarse una inadecuada clasificación de los sujetos efectuada exclusivamente a tenor de los resultados en una prueba de carácter estático.

Otro aspecto que nos parece sumamente interesante del análisis cualitativo efectuado por Resing, es el referente a la posibilidad de establecer las funciones cognitivas deficitarias en el alumnado participante mediante el análisis del tipo de sugerencias que estos requerían para poder resolver con éxito las tareas. La consideración de estas diferencias tiene importantes implicaciones, tanto en la determinación del potencial de aprendizaje como objetivo de la evaluación, como para la posterior intervención educativa. Así, se considera que los alumnos que requieren sólo algunas sugerencias de tipo metacognitivo para poder solucionar la tarea, tendrían en realidad pocos problemas para aplicar el razonamiento inductivo a los problemas propuestos, pero necesitarían indicaciones que, de alguna forma, «provocaran» la puesta en marcha de mecanismos o estrategias generales que centren, orienten, seleccionen o valoren las habilidades más específicas que deben aplicarse (Nisbett y Shucksmith, 1987); en cambio, alumnos que continuamente necesitan sugerencias más específicas y que incluso deben recibir algunas sugerencias cercanas a la solución correcta de la tarea, tendrían mayores dificultades en razonamiento inductivo. También es relevante la distinción entre alumnos que se benefician de las primeras sugerencias ofrecidas en la fase de entrenamiento, disminuyendo la demanda de éstas progresivamente, y alumnos que necesitan indicaciones de forma continuada hasta el final de las sesiones de entrenamiento. En el primero caso, el alumno mostraría la capacidad de razonar inductivamente cuando le muestran estrategias eficaces para hacerlo, beneficiándose claramente de la intervención; en cambio, el segundo de los casos mostraría quizás el más preocupante de los perfiles posibles que puede obtenerse con las pruebas dinámicas, es decir, una ausencia —o en todo caso un bajo nivel— de la capacidad de aprendizaje, en este caso, del razonamiento inductivo. De otra forma, lo que podría constituir una explicación alternativa a la anterior, los resultados mostrados en este segundo caso indicarían que los problemas propuestos en la prueba se hallan fuera de la ZDP de los sujetos.

Por último, también el análisis de las respuestas que justificaban las decisiones de los sujetos, fue considerado una fuente valiosa de información sobre el proceso de resolución de los problemas de razonamiento inductivo. Este análisis reveló, por ejemplo, que un número importante de alumnos que acertaban sus respuestas, no sabían por qué ésta era la correcta, o manifestaban grandes dificultades para justificar sus elecciones, ni tampoco cómo habían llegado a esa conclusión. Estos resultados se encuentran en línea con los hallazgos de otros muchos autores con relación a las dificultades observadas cuando se requiere el acceso de los procesos cognitivos a un nivel consciente, así como su verbalización, sobre todo en niños pequeños (Garner, 1987; Flavell, 1992; Garner y Alexander, 1989; Baker, 1989). Ya decíamos en páginas anteriores, cuando hacíamos referencia a la verbalización de los procesos de comprensión en el transcurso de una entrevista metacognitiva, que estas dificultades podrían tener lugar en función de la capacidad de reflexión metacognitiva del sujeto, pero que también podrían reflejar que los procesos puestos en práctica para resolver las tareas están muy automatizados. Esta distinción no parece estar efectuada en los análisis de Resing, quizás porque tratándose de alumnado de corta edad (7 y 8 años), la posibilidad de que los procesos cognitivos y metacognitivos puestos en práctica durante la resolución de las tareas estén automatizados, no parece viable. En este caso, parece lógico pensar que las dificultades encontradas se deban a los problemas generalizados que manifiestan los niños de estas edades para describir lo que ha ocurrido en sus mentes para que hayan podido solucionar un problema, así como para justificar el porqué de sus respuestas.

En definitiva, la prueba dinámica propuesta por Resing, la evaluación del potencial de aprendizaje a través del razonamiento inductivo (LIR), parece ser una contribución importante al campo de la investigación que nos ocupa y ha supuesto uno de los trabajos más interesantes desarrollados en consonancia con el método de sugerencias graduadas de Campione y Brown. Su preocupación por establecer algunas propiedades psicométricas sólidas y el análisis cualitativo adicional que, sin duda, enriquece las conclusiones que pueden deducirse de su aplicación, son «fortalezas» de esta propuesta. Sin embargo, como posibles debilidades, apuntadas ya por la propia autora (Resing, 2000), podríamos hacer alusión a aspectos nada banales como la necesidad de establecer más claramente la relevancia del razonamiento inductivo como elemento crucial en los procesos cognitivos, lo cual se sugiere en el trabajo desarrollado por Resing, y, más concretamente, el papel que juega este tipo de razonamiento en los procesos cognitivos infantiles. También es necesario determinar hasta qué punto son generalizables los resultados obtenidos en cuanto a la conducta manifestada por los niños estudiados en la solución de los problemas, ya que el hecho de que éstos fueran presentados para ser resueltos mediante razonamiento inductivo, lo que puede quedar alejado de muchas prácticas escolares, puede introducir serias dudas acerca de la extrapolación de conclusiones así obtenidas, al ámbito del rendimiento escolar.

Por otra parte, el contenido descontextualizado de sus dos tipos de tareas introduce para nosotros el principal problema a la hora de poder generalizar resultados; en este sentido cabe preguntarse sobre qué tipo de actividad escolar podríamos generalizar los mismos. Si se piensa que puede hacerse sobre cualquier tipo de actividad en la cual sea necesario aplicar el razonamiento inductivo, independientemente del dominio específico, habría que tener muy claro que este tipo de razonamiento subyace de forma general en nuestro procesamiento cognitivo una vez aprendido y que podría ser utilizado en situaciones de aprendizaje distintas a las que originaron su aprendizaje, pero esto no está investigado, o no lo hemos encontrado en los estudios de Resing. De otro modo, lo que a nuestro juicio sí podría ser un elemento que contribuyera a generalizar resultados indicativos de una mayor capacidad de aprendizaje, sería el componente metacognitivo de la metodología empleada en la fase de entrenamiento; en este sentido, cabría esperar, dada la naturaleza más reconocida de estos componentes como procesos generales que actúan a su vez sobre habilidades más específicas, que la interiorización de estrategias de control y regulación cognitiva empleadas por el evaluador en la fase de las sugerencias graduadas, pasen a formar parte de forma progresiva del repertorio metacognitivo del sujeto, permitiéndole actuar de forma igualmente estratégica en la resolución de otras tareas, independientemente de su contenido. Sobre este delicado pero interesante asunto volveremos con más insistencia cuando analicemos los resultados obtenidos por nosotros mismos en el estudio presentado.

La Prueba del Potencial de Aprendizaje de M. Budoff

Bases teóricas y metodológicas

La hipótesis de partida de este autor es similar a la formulada por Feuerstein y, en general, a la asumida por todos aquellos que han investigado las pruebas dinámicas; es decir, que algunos niños mentalmente deficientes o con desventaja sociocultural tenían una capacidad de aprendizaje distinta y mayor que la que reflejaban los resultados de pruebas estáticas de inteligencia; por tanto, si se les ofrecía la posibilidad de recibir una intervención educativa estructurada mediante unas instrucciones específicas, dirigida a

resolver un determinado problema, algunos de estos niños con dificultades verían incrementado su nivel de competencia y de respuestas adecuadas a las situaciones problemáticas planteadas de una forma superior a la que augurarían las pruebas convencionales (Budoff, 1968). Budoff compartía también con Feuerstein la idea de que muchos de los niños y niñas que, como consecuencia de un uso indebido de las pruebas estáticas de inteligencia, han sido valorados y clasificados incorrectamente por el sistema educativo, veían también infravaloradas sus capacidades reales al no tenerse en cuenta aspectos como su potencial de aprendizaje (Sternberg y Grigorenko, 2003). Por otra parte, en lo que a veces se convierte en un círculo vicioso difícil de romper, las condiciones socioculturales de partida que, como sabemos, condicionan al menos en parte las experiencias educativas enriquecedoras de las que se participa, o la presencia de déficits cognitivos, también influían negativamente en la valoración que se hacía de sus capacidades, básicamente porque esta valoración no tenía en cuenta la historia previa de los sujetos.

La población a la que se dirige la prueba desarrollada por Budoff es amplia y abarca, por una parte, a niños de bajo rendimiento, que pueden obtener bajas puntuaciones en pruebas de CI y que habitualmente pertenecen a minorías étnicas o a población privada socioculturalmente, y, por otra, a niños con un déficit cognitivo leve.

Tras la sintonía con el enfoque propuesto por Feuerstein, manifestada en los aspectos concernientes al marco de referencia, Budoff destacó una serie de inconvenientes que caracterizaban la propuesta contenida en el LPAD y que, en su mayor parte, tenían que ver con la débil estructuración de su fase de mediación y con los problemas metodológicos que conllevaba la ausencia de estandarización. De este modo, plantea la necesidad ineludible de normalizar la fase de intervención (Budoff, 1987a), por lo que en su propuesta se recogen sólo pruebas normalizadas, con una fiabilidad y validez contrastadas. El objeto de la fase de intervención de la prueba desarrollada por Budoff se dirigirá a la familiarización con las distintas pruebas empleadas, con el propósito de minimizar o neutralizar el efecto de la experiencia previa en los sujetos. Para estos fines, se han desarrollado versiones de aplicación dinámica basadas en varias pruebas normalizadas ya existentes, como las matrices progresivas de Raven o los cubos de Kohs. La aplicación de estas pruebas puede realizarse por separado o de forma combinada, y pueden ser administradas tanto individualmente como en grupo. Al contrario de lo que ocurría con el dispositivo de evaluación del potencial de aprendizaje de Feuerstein, estas pruebas que conforman el conjunto de las medidas del potencial de aprendizaje de Budoff, no suponen una gran dificultad en su aplicación y no exigen una formación prolongada (Budoff, 1987a).

En la fase de intervención, la actuación del evaluador se centra en (a) dirigir la atención del sujeto, (b) explicarle o aclararle los contenidos de la prueba y el procedimiento, y (c) guiarle en la ejecución de las tareas, a fin de encontrar la solución correcta. Por otra parte, cada prueba tiene un conjunto específico de instrucciones para el evaluador y unos métodos o principios más generales que orientan cada paso del proceso. Budoff señala estos principios básicos que debe seguir el evaluador para cada una de las tareas propuestas en la fase de intervención. A modo de ejemplo, para la tarea con los cubos de Kohs, estos principios serían los siguientes: (1) maximizar el éxito, señalando primeramente los elementos y contenidos más simples del diseño, (2) ofrecer elogios y ánimos frecuentes durante la realización de las actividades, (3) destacar la importancia de comprobar las construcciones de los bloques mediante la comparación con unas

tarjetas que muestran el diseño al sujeto, y (4) destacar el diseño de dos colores de los bloques. De esta forma, el evaluador realiza una intervención gradual en función de las dificultades que manifieste el sujeto. Es importante destacar en este punto que la estructuración y normalización de esta fase de instrucción no es del todo absoluta, sino más bien aproximada (Sternberg y Grigorenko, 2003). Posteriormente a esta fase de intervención tiene lugar la aplicación nuevamente de las pruebas que sirvieron para establecer las puntuaciones iniciales. Este postest se aplica habitualmente una segunda vez pasado un mes desde el final de la intervención.

La medida del potencial de aprendizaje se obtiene a raíz del procedimiento de instrucción e inicialmente permitió distinguir tres tipos de sujetos:

- *puntuadores altos*: obtienen altas puntuaciones previas a la intervención,
- *mejoradores*: obtienen bajas puntuaciones en el pretest, pero mejoran significativamente con la intervención, y
- *no mejoradores*: obtienen bajas puntuaciones en el pretest y no experimentan mejoras con la intervención.

Esta clasificación cualitativa (Budoff, 1987a; 1987b) recibió algunas críticas relativas a la dificultad que ofrecía para la distinción de casos o grupos específicos que requerían una mayor definición (Lidz, 1991), y fue modificada en trabajos posteriores. La nueva clasificación propuesta por Budoff no contemplaba la distinción cualitativa de los sujetos y estaba basada en conjuntos de puntuaciones operativizadas a raíz de los resultados obtenidos; de esta forma, se distinguía ahora entre: *puntuaciones preinstrucción*, *puntuaciones postinstrucción* y *puntuaciones postinstrucción ajustadas para el nivel pretest*, es decir la puntuación residual (Sternberg y Grigorenko, 2003).

Investigación y revisión crítica del modelo

A pesar de que inicialmente Budoff había rechazado la posibilidad de que el potencial de aprendizaje y el rendimiento escolar estuviesen relacionados, algunos de sus primeros estudios mostraron que las medidas de su prueba pronosticaban más certeramente el rendimiento en actividades escolares que las pruebas de CI, especialmente cuando se trataba de niños con necesidades educativas especiales (Budoff, Corman y Gimón, 1976). Budoff hizo referencia en numerosas ocasiones a la existencia de una alta correlación entre los resultados obtenidos con la prueba del potencial de aprendizaje y las impresiones o valoraciones que efectuaban los docentes (Budoff y Hamilton, 1976).

En uno de estos estudios se analizó la validez predictiva de la medida ofrecida por la prueba del potencial de aprendizaje en actividades de formación profesional (Budoff, Meskin y Harrison, 1971). Se establecieron tres grupos de alumnos en función de sus puntuaciones en la medida del potencial de aprendizaje. Los alumnos participantes recibieron una instrucción específica sobre electricidad, estableciéndose posteriormente la correlación entre la medida dinámica obtenida en esta tarea y una medida del rendimiento posterior. Los resultados mostraron que la medida del aprendizaje potencial predecía mejor las diferencias en el rendimiento en la prueba sobre electricidad que las medidas de CI.

En otra investigación, conducida por Laughon (1990), se analizó la validez predictiva de la Prueba del Potencial de Aprendizaje con respecto al rendimiento escolar, tomando esta vez como variable criterio la valoración que realizaban los educadores sobre el rendimiento de sus alumnos. El análisis de las correlaciones efectuadas entre las medidas posttest del potencial de aprendizaje y las valoraciones del rendimiento, se llevó a cabo en función de la diferenciación cualitativa que propuso originalmente Budoff; de esta forma, con base en las puntuaciones obtenidas en la medida posttest de la prueba se distinguieron los tres grupos, es decir: *puntuadores altos* —que obtuvieron puntuaciones superiores a la media—, *mejoradores* —que obtuvieron un nivel intermedio— y *no mejoradores* —con puntuaciones inferiores a la media—. Los resultados mostraron que las correlaciones eran muy superiores a las que ofrecían CI y rendimiento escolar para los alumnos de nivel intermedio (mejoradores) e inferior (no mejoradores); sin embargo, eran similares a las ofrecidas por el CI y el rendimiento en los alumnos que puntuaban por encima de la media.

Un aspecto sin duda interesante de los estudios de Budoff y sus colaboradores fue el intento de establecer relaciones entre las medidas del potencial de aprendizaje y una serie de variables cognitivas, emocionales y demográficas. En uno de esos estudios (Budoff, 1987a), se evaluaron los perfiles de personalidad de sujetos con un bajo CI que habían puntuado en la prueba del potencial de aprendizaje por encima de la media, y se compararon con los de los sujetos restantes del mismo grupo. Los resultados indicaban que los sujetos con un mayor potencial de aprendizaje mostraban indicadores mayores de motivación, flexibilidad cognitiva y amabilidad, así como un menor grado de impulsividad. En este mismo estudio se correlacionaron las medidas del potencial de aprendizaje con indicadores socioeconómicos, hallándose una relación entre la cantidad de intervención necesaria para la mejora y la condición socioeconómica de los sujetos; al mismo tiempo, Budoff destacó que estos alumnos se beneficiaban más de una instrucción individual que grupal en comparación con el alumnado que no presentaba desventaja social y económica.

En otros estudios de corte longitudinal (Budoff, 1967; 1987a; Budoff y Friedman, 1964) se hizo un seguimiento continuado a largo plazo a un grupo de chicos con CI bajo, a los que se aplicó la clasificación cualitativa referida anteriormente, en función de sus puntuaciones. Se concluyó que la categoría correspondiente a los *mejoradores*, en comparación con los *no mejoradores*, estaba compuesta por chicos que tenían mejor pronóstico en una serie de factores relacionados con su vida escolar y social, como por ejemplo: rendir mejor en programas de educación especial, alcanzar mayor independencia económica y social como adultos, tener amigos y relaciones de pareja, u obtener el permiso de conducir.

El análisis reiterado que efectúa Budoff de las relaciones que mantiene el potencial de aprendizaje con características individuales o socioeconómicas, sitúa a éste como una característica individual más, un atributo personal que puede estar relacionado con un mayor ajuste social o con rasgos de personalidad más adaptativa; sin embargo, como destacan Sternberg y Grigorenko (2003), a pesar del interés que estos estudios pueden despertar, aspectos como la metodología correlacional utilizada, la heterogeneidad de las muestras, así como la ausencia de unos controles adecuados, aconsejan mucha cautela en la interpretación y generalización de algunas de las conclusiones efectuadas desde este enfoque.

Otro problema metodológico, planteado a raíz de la conversión de las categorías cualitativas formuladas en función de la distribución de las puntuaciones obtenidas (puntuadores altos, mejoradores y no mejoradores) en un conjunto de puntuaciones continuas distribuidas cuantitativamente (preinstrucción, postinstrucción y puntuación residual), es la necesidad de establecer unos marcos de referencia estándar para las distintas poblaciones estudiadas, que permitan establecer comparaciones de las puntuaciones obtenidas con puntuaciones normativas (Sternberg y Grigorenko, 2003). En este sentido, aunque estos últimos autores citados, y también Lidz (1991), destacan ciertas ventajas de la operativización de las distinciones efectuadas en los grupos de sujetos, parece que la utilización de la descripción cualitativa en los estudios presentados por Budoff introduce un elemento novedoso que permite una interpretación enriquecida de los resultados, por cuanto informa del comportamiento de los sujetos con relación al aspecto esencial en la ED, es decir, la interacción mediada en la fase de intervención.

Como conclusión a la revisión realizada en torno a las aportaciones del enfoque de Budoff al estudio de la ED, podemos decir, de acuerdo con Sternberg y Grigorenko, que la Prueba del Potencial de Aprendizaje ha logrado contribuir a una diferenciación efectiva entre niños que, obteniendo puntuaciones bajas en pruebas estáticas de CI, muestran un comportamiento y unos resultados muy distintos cuando son instruidos específicamente en el conocimiento y manejo de esas mismas pruebas. La estrategia dinámica propuesta por Budoff se ha mostrado especialmente adecuada para este objetivo concreto, estableciendo, así mismo, diferencias entre las categorías propuestas de alumnos mejoradores y no mejoradores en cuanto a la potencia del nivel de predicción de la prueba con respecto al rendimiento escolar y, en mayor medida, con respecto a factores cognitivos y de adaptación socio-personal.

Por último, una reflexión más que tiene que ver con la utilización de pruebas estáticas y no contextualizadas en ausencia de un marco teórico de referencia. Como dijimos, anteriormente, Budoff y sus colaboradores versionaron dinámicamente algunos tests cognitivos. Esta introducción de pruebas estáticas supuso una novedad y, como recordamos, estuvo originada en la búsqueda de una mayor fiabilidad y validez; sin embargo, el contenido de estas pruebas, o al menos de las más comúnmente utilizadas, tiene una escasa relación con actividades escolares y sociales significativas para los sujetos que las realizan, habitualmente escolares. Al igual que ocurría en el enfoque de Feuerstein o con las aportaciones de Tzuriel, las medidas utilizadas para determinar el potencial de aprendizaje no tienen en cuenta los contenidos escolares ni el contexto escolar, adonde, por otra parte, se dirigen sus propuestas principalmente. De esta forma, se aportan estudios y resultados sobre predicción del rendimiento escolar, pero la conexión entre las actividades de evaluación y las actividades de enseñanza y aprendizaje que tienen lugar habitualmente en la escuela, no parece clara. Por otra parte, el enfoque de Budoff, tal y como apuntábamos antes, no parece estar sustentado por un marco teórico de referencia (Saldaña, 2001; Sternberg y Grigorenko, 2003), lo que origina que los resultados obtenidos deban ser interpretados con base en otros postulados teóricos. En relación con esto, la propuesta de Budoff no aporta tampoco un conjunto de funciones o procesos cognitivos o sociopersonales que sean el objeto de la modificación —como sí ocurría en modelos revisados anteriormente—, por lo que a nuestro juicio, esto dificulta la necesaria conexión entre la interpretación de la evaluación realizada y una posible propuesta de intervención posterior para la mejora de aquellos aspectos encontrados deficitarios.

La Evaluación del Potencial de Aprendizaje (EPA), de R. Fernández-Ballesteros

Bases teóricas y metodológicas

El trabajo desarrollado en nuestro contexto por Fernández-Ballesteros y sus colaboradores (Fernández-Ballesteros, 1996; Fernández-Ballesteros y Calero, 1993, 2000; Fernández-Ballesteros, Espinosa, Colom y Calero, 1997) se halla influenciado notablemente por el anteriormente analizado de Budoff. Esta influencia está marcada básicamente por tres razones: (a) la utilización adaptada de los materiales que a su vez adaptara y utilizara Budoff en su prueba del potencial de aprendizaje, (b) el uso de un paradigma cuantitativo en el análisis de la información obtenida, y (c) la asunción del razonamiento esgrimido por Budoff relativo a la idoneidad de entrenar en el manejo y conocimiento de pruebas de aptitud normalizadas, como las matrices progresivas de Raven, considerando que éste es el mejor modo de minimizar la influencia del aprendizaje previo del sujeto y pudiendo valorar así la modificabilidad experimentada a través de la intervención. Enmarcado en estas ideas se desarrolló un primer instrumento de Evaluación del Potencial de Aprendizaje (EPA), (Fernández-Ballesteros, Calero, Campllonch y Belchi, 1987/1990), que ha sido posteriormente revisado y ampliado en los últimos años por sus mismos autores (Fernández-Ballesteros et. al., 2000). Se trata de una prueba que puede ser aplicada de forma individual y colectiva a partir de 5 años. Consiste básicamente en la aplicación del test de matrices progresivas de Raven en las fases pretest y postest de la evaluación, mientras que durante la fase de intervención se trabaja con un total de 132 diapositivas con las que se trata de profundizar en el conocimiento y uso de las estrategias necesarias para resolver el test.

Mediante este formato o diseño de evaluación clásico pretest/ mediación /postest, en el EPA se tratan de establecer los cambios entre una «línea base», otorgada por la puntuación pretest, obtenida mediante la pasación de la prueba en su forma convencional, y el nivel alcanzado una vez finalizada la fase de entrenamiento, obtenido mediante la puntuación en la fase postest. El análisis de esta diferencia entre ambas puntuaciones nos daría una medida del potencial de aprendizaje. La propuesta de Fernández-Ballesteros también asume, como decíamos un poco más arriba, la utilización de los materiales usados en el enfoque de Budoff; el uso de estos materiales, especialmente de las matrices progresivas, responde a una concepción de la ED defendida también por Feuerstein y otros investigadores, en la que los instrumentos de evaluación mediante los que establecer el potencial de aprendizaje debían tener el mínimo contenido posible de tipo verbal, así como la menor relación posible con contenidos culturales. Esta idea, junto al progresivo conocimiento que se ha ido teniendo de las estrategias necesarias para resolver cada uno de los ítems del test de Raven (Hunt, 1974), facilitando de esta forma el desarrollo de programas de entrenamiento bien fundamentados (Powell, Love y Vega, 1976; Johnson, 1978), ha contribuido decisivamente a la extensión de su utilización en pruebas dinámicas. Otro aspecto importante en este sentido es la alta fiabilidad test-retest que presenta la aplicación de esta prueba (en torno a 0,90); este hecho, según Fernández-Ballesteros y Calero (2000), aconseja especialmente su uso en una prueba de ED, ya que de esta forma, los resultados de la fase de entrenamiento se tornan decisivos para determinar los cambios observados entre la ejecución pre y postentrenamiento. Se argumenta también por parte de la autora de la prueba y sus colaboradores, que el bajo contenido verbal de esta prueba la hace especialmente idónea para la evaluación de niños en situación de

desventaja sociocultural; al igual que Feuerstein o Budoff, las razones aportadas se relacionan con que las mínimas influencias de tipo cultural en el test de Raven igualarían o controlarían el efecto de la historia previa del sujeto y otorgaría la posibilidad de establecer el potencial de aprendizaje de forma independiente a las condiciones de partida del sujeto (Feuerstein, Rand y Hoffman, 1979; Budoff y Hutten, 1971; Carlson y Wiedl, 1979; Kaniel y Tzuriel, 1992).

Siguiendo los criterios fundamentales que sirven a Fernández Ballesteros y sus colaboradores para clasificar las distintas propuestas de ED, el EPA podría ser definido como un instrumento altamente estructurado; según la propia autora, esta estructuración vendría dada a su vez en función de tres criterios: (a) por el énfasis en el producto final del entrenamiento —número de participantes que han obtenido ganancias—, (b) por el tipo de interacción mantenida en la fase de intervención, interacción diseñada previamente y estandarizada, y (c) por el objetivo final de la evaluación, en este caso la valoración o medida del potencial de aprendizaje. Esta descripción se realiza por oposición a un modelo de ED más clínico o cualitativo (p.e. el modelo de Feuerstein), en el que el interés primordial del evaluador es el proceso mismo de aprendizaje, la interacción mantenida se caracteriza por su flexibilidad y adaptación al sujeto evaluado, y el objetivo de la evaluación es el establecimiento de las funciones cognitivas deficientes que manifiesta el sujeto en sus intentos de resolver las tareas. Aunque, bajo nuestro punto de vista, la clasificación utilizada no parece ser muy inclusiva con respecto a la gama de propuestas de ED surgidas en los últimos años, y más bien parece constreñirse al contraste entre el modelo propuesto por Feuerstein y el enfoque alternativo que sugirió Budoff, esta descripción y diferenciación nos sirve para encuadrar mejor la propuesta recogida en el EPA.

La estructuración a la que hacíamos referencia se concreta en la fase de intervención de la prueba, de modo que ésta ofrece unas pautas para el diálogo que lo orientan al proceso de aprendizaje y a la generalización (una muestra de la estructura del entrenamiento para uno de los ítems puede observarse en el cuadro 17). La intervención (a) está basada en principios como el modelado o el reforzamiento, (b) incorpora también la retroalimentación inmediata al sujeto sobre su actuación, (c) trata de provocar la verbalización de las respuestas y de las explicaciones que las justifiquen, así como de (d) analizar las estrategias seguidas en cada ítem para poder dar solución a los mismos. Por otra parte, cada situación problema del instrumento es trabajado mediante la proyección de dos diapositivas; en una primera se presenta el estímulo y en la segunda se aborda la respuesta solución. Existen dos niveles de aplicación en función de las puntuaciones obtenidas por los sujetos en la evaluación pretest. El nivel I aborda 39 posibles problemas mediante 74 diapositivas y entrena las series A, Ab y B del test de Raven; el nivel II aborda 29 problemas con 58 diapositivas, entrenando las series C, D y E del test. Se obtienen tres puntuaciones: pretest, postest y puntuación de mejora; esta puntuación de mejora sirve a los autores para establecer la clasificación cualitativa ya propuesta por Budoff de *puntuadores altos*, *mejoradores* y *no mejoradores*.

La aplicación del EPA permite también un análisis cualitativo en función de cuatro categorías: (1) los errores cometidos por los sujetos durante su administración, (2) el análisis de la información, (3) la modalidad de respuesta, y (4) la apropiación de la respuesta. En cuanto a la primera de estas categorías, se toman en consideración los tipos de errores que más frecuentemente se hayan cometido en la realización de las

*Cuadro 17. Ejemplo de la fase de entrenamiento en el EPA (ítem 24 del test de Raven).
Tomado de Fernández-Ballesteros y Calero, 2000*

Ítem 24

El evaluador dice:

...y al mismo tiempo...

- | | |
|--|--|
| - Este es un tipo diferente de puzzle, pero es divertido. | Se muestra la diapositiva estímulo |
| - Vamos a averiguar que pieza va aquí. | Se indica el lugar vacío |
| - Aquí podemos ver algunos cuadrados y círculos, y dentro algunas líneas y puntos. | Se indica |
| - Podemos mirar de arriba abajo y de izquierda a derecha. | Se dibuja una flecha para mostrar esto |
| - En este puzzle algo cambia en esta dirección. | Se indica la dirección horizontal |
| - Aquí tenemos una regla para hacer estos puzzles. Lo que cambia cuando tú vas de arriba abajo en la primera columna, también tiene que cambiar en la otra. | Se indica la primera columna
Se indica la segunda columna
Se indican las figuras |
| - Entonces aquí, el cuadrado cambia a círculo cuando tú vas de arriba abajo. | |
| - Si esto pasa aquí (en la primera columna) el cuadrado también tiene que cambiar a un círculo aquí. Sabemos que la figura correcta es un círculo. | Se indica
Se dibuja |
| - ¿Qué tiene que ir dentro del círculo? | Se indica |
| - Tenemos que mirar la fila horizontal para ver qué cambia. Las líneas cambian a puntos. Lo que cambia aquí de izquierda a derecha debería cambiar aquí también. | Se indica la primera fila
Se indica la segunda fila |
| - Entonces, dentro del círculo deberían ir... puntos. | Se muestra la diapositiva con la solución |
-

tareas, bien al inicio, antes de comenzar el entrenamiento, o bien los errores que persisten durante la intervención. El análisis se realiza siguiendo los trabajos de Foulds (1949) y de Foulds y Raven (1950), citados en Fernández-Ballesteros y Calero (2000), sobre el contenido de la prueba y se aplica en función de una serie de parámetros: (a) establecimiento de relaciones incompletas, que supone una elección parcial de las relaciones correctas sin tener en cuenta todas las variables relevantes del problema, (b) principios de razonamiento erróneos, que supone la verbalización de fallos de razonamiento que no permiten solucionar adecuadamente el problema, (c) confluencia de ideas, que hace alusión al hecho de tener en cuenta aspectos irrelevantes de la figura presentada (p.e. dar una respuesta determinada en función de su atractivo), (d) repetición, que supone la elección de una alternativa similar a una parte del dibujo presentado, y (e) persistencia, que consiste en la selección de la alternativa de respuesta que ocupa la misma posición en la hoja que la elección anterior. Por su parte, el análisis de la información hace referencia a la anotación de cómo procesa el sujeto la información proporcionada mediante las diapositivas; este análisis se efectúa tomando en consideración dos posibles estilos cognitivos o conductuales: impulsividad o no impulsividad. La tercera de las categorías, la modalidad de respuesta, implica la anotación del tipo de respuesta que el sujeto da de forma espontánea; ésta se clasifica a su vez en función de cuatro categorías: gráfica, gestual, verbal y anticipatoria (los

sujetos describen las partes o elementos del problema, pero no son capaces de dar la respuesta correcta). Por último, la apropiación de la respuesta implica la evaluación del modo en que se produce esta apropiación por parte del sujeto; para ello se utilizan varias categorías: (a) respuesta correcta dada de forma espontánea, (b) corrección espontánea del sujeto, (c) corrección después de la retroalimentación dada por el evaluador, (d) corrección después del entrenamiento, que supone la revisión de una respuesta con posterioridad a una secuencia completa de entrenamiento de un ítem, y (e) respuesta incorrecta.

Investigación y revisión crítica del modelo

En el cuadro 18 podemos observar una síntesis cronológica de los estudios realizados por Fernández-Ballesteros y sus colaboradores. Un buen número de estos estudios ha tenido entre sus principales objetivos la valoración de distintos aspectos psicométricos de la prueba. De esta forma, algunos se han dirigido principalmente a determinar la *efectividad del tratamiento*, entendida como la capacidad para generar cambios en la actuación del sujeto que no sean debidos a otros factores ajenos al propio tratamiento y que no se deban a la simple repetición de la prueba de matrices; también han tratado de valorar la *fiabilidad*, entendida como estabilidad de los resultados obtenidos por diferentes grupos de población y, por último, los estudios se han centrado en la *validez predictiva* sobre el cambio en el funcionamiento cognitivo del sujeto medido durante un periodo más largo de aprendizaje. Estos tres aspectos han sido evaluados con diferentes segmentos de población, diferenciados en función de sus puntuaciones en pruebas normalizadas de inteligencia, y utilizando grupos de control.

En cuanto a la efectividad del tratamiento, los estudios de Campllonch y Fernández-Ballesteros (1981) y Fernández-Ballesteros, Campllonch y Macia (1982), demostraron que los cambios producidos después del entrenamiento en sujetos de 12-15 años con distintos nivel intelectual eran independientes de su nivel inicial en el pretest. Se han confirmado estos estudios posteriormente en un segmento de edad similar y bajo funcionamiento cognitivo, contando con un grupo control y estableciendo diferencias significativas en los cambios producidos (Fernández-Ballesteros y Calero, 1993) y también con ancianos (59-87 años) (Fernández-Ballesteros y Calero, 1995). Algunos de estos mismos estudios sirvieron para evaluar la fiabilidad del cambio producido mediante la valoración de la estabilidad de las ganancias después de un tiempo. En el estudio citado de Fernández-Ballesteros y Calero (1993), las ganancias postest se mantuvieron 6 meses después del entrenamiento, con diferencias significativas entre los grupos experimental y control; también se obtuvieron similares resultados en el otro de los estudios citados de estas dos autoras, realizado con población mayor. En cuanto a la validez predictiva de la prueba con respecto a otras medidas del aprendizaje del sujeto, efectuadas tras periodos de entrenamiento más prolongados, ésta fue evaluada con relación a los efectos que sobre una prueba estática (WISC) tenía la participación de los sujetos en el Programa de Enriquecimiento Instrumental (PEI) de Feuerstein. En el estudio (Fernández-Ballesteros y Calero, 1993) participaron 48 alumnos, 24 fueron incluidos en el GE y recibieron el PEI durante 9 meses, los otros 24 alumnos conformaron el GC; las puntuaciones iniciales del total de los sujetos en el WISC oscilaban entre 55 y 80. Se les evaluó dinámicamente mediante el EPA con el objetivo apuntado de predecir las posibles mejoras obtenidas tras su participación en el PEI. Los resultados mostraron que las puntuaciones de mejora obtenidas en el EPA podían predecir de forma significativa las ganancias obtenidas por los sujetos en la aptitud

verbal, medida con el WISC; por otra parte, la clasificación cualitativa habitualmente establecida en función de las puntuaciones de mejora en el EPA entre mejoradores, o ganadores, y no mejoradores, servía para predecir las ganancias obtenidas tras la aplicación del PEI, de forma que los ganadores mejoraron significativamente más que los alumnos clasificados como no mejoradores.

Cuadro 18. Investigaciones sobre el EPA. Tomado de Fernández-Ballesteros y Calero (2000)

Autores	Fecha	N	Edad	Contexto	Características
Campllonch y Fernández-Ballesteros	1981	90	12-15	Escuela pública	Deprivación sociocultural
Fernández-Ballesteros, Campllonch y Macia	1982				CI (WISC) 70-115
Calero y Belchi	1988	64	12-15	Escuela pública	CI (Cattell) 80-110
	1988	66	11-15	Escuela rural	CI (Cattell) 80-110
Fernández-Ballesteros y Calero	1993	48	10-14	Centro de E. Especial	Deprivación sociocultural
	1995				CI (WISC) 50-80 —24 “déficit orgánico”—
Calero	1994	156	13-16	Escuela pública	Deprivación sociocultural (desfase curricular de al menos 2 años) CI 90-110
Calero y Lozano	1994	21	60-80	Centros de día	Nivel educativo bajo
Fernández-Ballesteros y Calero	1995	90	58-83	Centros de día y Residencias	Nivel educativo bajo
Calero y G- Berben	1997	25	60-85	Centros de día y Residencias	96% con menos de 4 años de escolarización

Otros estudios del grupo de Fernández-Ballesteros se han preocupado por la distinción entre mejoradores y no mejoradores para explicar la heterogeneidad de la población con déficit intelectual (Sternberg y Grigorenko, 2003). Se investigaron las diferencias en las mejoras obtenidas por sujetos con déficit cognitivo atribuible a causas manifiestamente orgánicas, y sujetos que no tenían causa conocida de sus deficiencias en el aprendizaje. Los resultados mostraron que las mejoras obtenidas por el alumnado sin déficit orgánico manifiesto eran el doble y presentaban una mayor estabilidad que las mejoras obtenidas por el alumnado con déficit cognitivo con causa orgánica. Sobre esta diferenciación entre alumnado con discapacidad cognitiva y alumnado que presentaría desventaja sociocultural o dificultades de aprendizaje, volveremos más adelante cuando hablemos del estudio presentado por nosotros mismos en este trabajo.

En una serie de estudios muy interesantes dirigidos a evaluar la transferencia de los resultados obtenidos a través del EPA hacia otras tareas con distinto grado de similitud, Fernández-Ballesteros y sus colaboradores se encontraron diferentes resultados. En sus primeros estudios (Campllonch y Fernández-Ballesteros, 1981; Fernández-Ballesteros, Calero, Campllonch y Belchi, 1987/1990), utilizando una primera versión del instrumento, estos investigadores llegaron a la conclusión de que los efectos producidos por el entrenamiento en el EPA tenían una transferencia, que en la distinción efectuada por Brown y sus colaboradores podría ser catalogada como «cercana», apareciendo en

tareas similares a las entrenadas o que requirieran la misma estrategia para ser resueltas (Brown y Campione, 1984; Campione, Brown, Ferrara, Jones y Steinberg, 1985). Estos estudios habían empleado, por una parte, el test de Dominós (TIG-1) (TEA, 1976) y el factor espacial del PMA (Yela, 1967; TEA, 1971), no encontrando diferencias entre los grupos experimental y control en la ejecución de estas pruebas tras haberles sido aplicado el EPA al GE; por otra parte, utilizando una versión mejorada del instrumento, habían tratado de evaluar la transferencia entre la ejecución en distintas series del test utilizado en la prueba, es decir, las matrices progresivas de Raven. No encontraron tampoco en esta ocasión resultados positivos, incluso cuando se evaluó la transferencia desde las series más complejas y difíciles (series C y E) a las series más fáciles (A, Ab, B y C). Una de las conclusiones más relevantes que estos investigadores extrajeron con relación a la transferencia tras estos estudios, era que el entrenamiento de estrategias cognitivas parecía caracterizarse por la especificidad de las tareas/estrategias implicadas. Tras esta serie de estudios, Fernández-Ballesteros y Calero (1995), interesadas en los procesos de transferencia y generalización de los resultados obtenidos con la aplicación del EPA más allá de la actuación de tareas muy similares, emprendieron una nueva investigación utilizando en este caso el test de factor «G» de Cattell (Cattell y Cattell, 1974) como tarea a través de la cual medir la transferencia; este test incluye un subtest de matrices y evalúa también el razonamiento inductivo. La muestra del estudio estaba compuesta por personas mayores, en su mayoría sin estudios primarios finalizados. Los resultados esta vez mostraron diferencias significativas entre el GE y el GC en cuanto a su desempeño, tanto en la ejecución postentrenamiento del test de Raven, como en el test de Cattell. Estos resultados fueron replicados y confirmados en un estudio de seguimiento tres meses después. La conclusión de las autoras sobre la eficacia del EPA para generar transferencia del aprendizaje apuntaba un poco más alto tras este estudio, haciendo referencia a que la transferencia del efecto del entrenamiento en la aplicación de la prueba podía producirse a otras tareas que midieran —utilizando nuevamente la diferenciación que sobre los posibles niveles de la transferencia se realizó en los estudios de Brown y sus colaboradores— el «mismo constructo» (factor G y razonamiento inductivo, en este caso) y requirieran estrategias similares a las que fueron entrenadas.

Quisiéramos adentrarnos ahora en una valoración un poco más profunda y crítica acerca del trabajo desarrollado por Fernández-Ballesteros y su equipo. En este sentido, debemos comenzar destacando el especial interés que nos ha suscitado el análisis cualitativo propuesto como posible alternativa o complemento a la valoración cuantitativa de la prueba desarrollada, contribuyendo, bajo nuestro punto de vista, a enriquecerla notablemente. Este análisis cualitativo permite ofrecer un perfil del sujeto a través del cual puedan destacarse aspectos relativos al tipo de errores que comete más frecuentemente en las tareas, a la modalidad de respuesta que tiende a emplear habitualmente, al grado de impulsividad que manifiesta en su actividad cognitiva y conductual, así como al modo y el momento en que el sujeto resuelve el problema. Sin duda, este encuadre de la propuesta de Fernández-Ballesteros parece ofrecer una información más valiosa para el contexto escolar, motivo por el cual hemos querido destacarlo. En este sentido, un análisis de este tipo proporcionaría información que puede ser utilizada en una posterior intervención educativa, si bien es cierto que los procesos cognitivos evaluados cualitativamente son puestos en práctica sobre contenidos que no se relacionan directamente con el contexto escolar ni con las actividades que allí se desarrollan habitualmente, lo que sigue manteniendo, bajo nuestro punto de vista, los problemas para poder efectuar una mayor generalización de

los resultados sobre actividades escolares con contenidos más específicos, permitiendo de esta forma planificar una intervención que conjugue ambas cosas: intervención sobre el funcionamiento cognitivo incardinada en actividades contextualizadas.

Sin duda, otro de los aspectos más interesantes de la propuesta y los estudios de Fernández-Ballesteros es la alusión al efecto de la intervención en la capacidad de autorregulación del sujeto. Esta idea, apuntada ya por Vygotski y desarrollada en parte por Feuerstein, forma parte de las que arguye esta autora en la explicación de los cambios producidos y, especialmente, de los cambios observados en la realización de actividades sin aparente relación con los contenidos del test entrenado. En efecto, Fernández-Ballesteros conocía los resultados de Budoff al estudiar la validez concurrente sobre el rendimiento escolar de su prueba del potencial de aprendizaje; las medidas postinstrucción obtenidas estaban relacionadas con la actuación del sujeto en pruebas de aptitud a excepción de aquellas cuyo contenido era fundamentalmente de tipo verbal. Este resultado no sorprendió a Budoff, en su convencimiento —sobre todo al inicio de sus investigaciones— de que el potencial de aprendizaje poco tenía que ver con tareas saturadas culturalmente, en cambio sí atrajo la atención de la psicóloga española y de sus colegas; en algunos de sus estudios (Fernández-Ballesteros y Calero, 1993, 2000; Fernández-Ballesteros, 1996; Fernández-Ballesteros et. al., 1997) se había encontrado con el dato de que las medidas del potencial de aprendizaje obtenidas con el EPA predecían precisamente la mejora de aptitudes verbales, aunque no la del rendimiento de los sujetos; así mismo, las puntuaciones del EPA predecían un rendimiento cognitivo superior para aquellos sujetos que habían mejorado las mismas durante la aplicación dinámica. Como apuntábamos anteriormente, la explicación sostenida por Fernández-Ballesteros y su equipo apunta al papel relevante que jugaría la interacción verbal que mantienen evaluador y alumno en la fase de mediación; esta interacción podría tener como contenido fundamental en las pruebas de ED la enseñanza del *control verbal ejecutivo* (Campione y Brown, 1979) y, por tanto, de autorregulación. Este hecho tendría lugar de forma independiente a que el contenido de los materiales entrenados fuera no verbal (las matrices progresivas). Este aspecto es especialmente relevante para nosotros, como veremos más adelante en nuestro propio estudio; la explicación de Fernández-Ballesteros y de sus colaboradores de los resultados obtenidos no hace más que reforzar la hipótesis de que una intervención basada en principios o estrategias metacognitivas posibilitaría en mayor medida la transferencia y generalización de los aprendizajes a situaciones o tareas más alejadas del contexto concreto donde fueron aprendidas (Campione y Brown, 1987; Campione, Brown y Ferrara, 1987; Mora, 1998; Navarro, 2004). Este entrenamiento metacognitivo tal vez no estuviera en la intención de los evaluadores de los estudios de Fernández-Ballesteros y sus colaboradores, pero parece que podría ser algo consustancial a los modelos de ED, ya que en definitiva se trata de *enseñar a aprender*, o dicho de otro modo, de un modelo de evaluación que se relaciona bien con las propuestas metodológicas basadas en el *aprender a aprender* (p.e. Nisbett y Shucksmith, 1987), en el que tanta relevancia es otorgada a los procesos metacognitivos.

Por otra parte, destacábamos en nuestra revisión de las bases teóricas de este modelo, que Fernández-Ballesteros, apoyada por numerosos trabajos, justificaba la utilización de las matrices progresivas de Raven en el EPA; sin embargo, queríamos introducir una reflexión acerca de algunas de estas justificaciones esgrimidas. En primer lugar, la descontextualización del contenido de la prueba utilizada —el test de Raven— puede influir notablemente en la generalización que podamos efectuar de los resultados

obtenidos. Incluso cuando estamos valorando elementos como el razonamiento inductivo, las posibilidades de que esta aptitud sea aplicada a tareas distintas de aquellas en las que fue entrenada, no están definidas ni evaluadas. La evaluación del potencial de aprendizaje ha sido propuesta para su uso en ámbitos clínicos y educativos; en este sentido, especialmente en este último, parecen tener una enorme importancia los contenidos escolares específicos en la aplicación de determinadas estrategias para la mejora del aprendizaje (Gardner, 1983; Chi, 1987; Carey y Spelke, 1994; Martín, 1999; Jensen, 2000; Mateos, 2001; Navarro, 2004). En segundo lugar, el hecho de que la ED base su razón de ser en el convencimiento de que muchos alumnos con bajas puntuaciones en pruebas de aptitud estáticas, se benefician de un proceso de evaluación que crea una situación de interacción educativa para valorar su capacidad de aprendizaje, teniendo en cuenta la historia educativa del sujeto y el contexto sociocultural, no implicaría bajo nuestro punto de vista la necesidad de utilización de instrumentos de evaluación con contenidos descontextualizados, con la pretensión de ser ajenos al contexto social y cultural donde tienen lugar los aprendizajes del sujeto. En la evaluación de la capacidad de aprendizaje de un sujeto, mediante tareas realizadas con la mediación de un adulto o compañero más experto, situándose en su ZDP, no parece imprescindible que estas tareas deban tener un contenido descontextualizado. Sin embargo, esta última consideración nos conduce al problema de que ante una medición de carácter cuantitativo, se requiere el control lo más exhaustivo posible de la variable conocimientos y experiencias previas del sujeto; por tanto, no obviamos las dificultades que conllevaría una propuesta de ED mediante instrumentos que contaran con contenidos contextualizados y que a su vez permitiera un análisis cuantitativo de los resultados. Estas dificultades —haciendo referencia a un formato pretest/ mediación /postest—podrían ser minimizadas mediante (a) una determinación exhaustiva previa del conocimiento de los sujetos sobre el contenido que se trate —habitualmente la pasación de una prueba de aptitud a modo de pretest podría ofrecernos esta información—, que nos permitiera valorar los resultados del cambio en función principal o exclusivamente de la fase de mediación, (b) un proceso de mediación dirigido especialmente a facilitar la puesta en práctica y el aprendizaje de aquellos procesos cognitivos que están implicados de forma significativa en el desarrollo y dominio de las actividades planteadas y muy especialmente de los procesos metacognitivos, (c) una medida del cambio producido con respecto a la línea base establecida; información que podríamos obtener a través de un postest equivalente al pretest, y (d) contando con grupos de control adecuados, que nos permitan validar unas conclusiones positivas en cuanto a la modificabilidad de los sujetos, en el caso de que queramos establecer comparaciones con otros sujetos a efectos de investigación o de otros propósitos.

Asumiendo un paradigma más cualitativo, no serían totalmente necesarias todas estas medidas, ya que en definitiva se trataría de analizar el proceso de aprendizaje de un sujeto en una situación de colaboración mediante el desarrollo de actividades que se sitúan en su ZDP con el fin de determinar el desarrollo potencial de su aprendizaje, siempre dinámico por otra parte, y establecer las medidas de apoyo necesarias para optimizar el proceso de enseñanza y aprendizaje, favoreciendo experiencias de aprendizaje mediado que sean significativas para el sujeto y, así mismo, el acceso progresivo de éste a formas de funcionamiento cognitivo más autónomas.

En definitiva, y concluyendo esta revisión efectuada sobre una pequeña parte del inmenso trabajo desarrollado por Fernández-Ballesteros, podríamos decir que la

investigación sobre la evaluación dinámica del potencial de aprendizaje que ha utilizado la prueba EPA, ha permitido revisar y perfilar paulatinamente el instrumento, de forma que se perfeccionaran sus propiedades psicométricas, como la fiabilidad del cambio en las puntuaciones de ganancia o la validez predictiva sobre otras medidas del constructo que es objeto de entrenamiento en la fase de intervención. Sin duda, esta preocupación por propiedades como la fiabilidad, la validez, así como por la generabilidad de los resultados obtenidos con su aplicación, unido al interés manifestado en sus estudios por apoyar empíricamente la heterogeneidad de la población mentalmente deficiente, o por la aplicación de procesos de evaluación dinámica a segmentos de población no comúnmente beneficiaria de estos trabajos, como personas mayores y también enfermos mentales (Fernández-Ballesteros y Calero, 2000), constituye a nuestro juicio una «fortaleza» importante de su propuesta, al tiempo que nos hace estar expectantes con respecto a los nuevos resultados que puedan derivarse de los estudios en marcha, así como a la posible evolución de estos últimos, quizás guiada por las interesantes conclusiones de sus investigaciones sobre la transferencia.

Las pruebas de aprendizaje de J. Guthke

Bases teóricas y metodológicas

Desde esta perspectiva de la ED se destaca la relevancia para el proceso educativo de evaluar la *capacidad de aprendizaje* de un sujeto, en contraste con la valoración de productos de aprendizaje ya adquiridos (Grigorenko y Sternberg, 1998). De este modo, se hace hincapié en la necesidad de incluir una sesión de aprendizaje en las pruebas de aptitud para remediar el hecho de que la información que puede obtenerse principalmente de estas pruebas, sea una información que hace referencia a habilidades ya «cristalizadas» o, simplemente, a la ausencia de estas habilidades. Para esto, en sesiones de aprendizaje con mayor o menor duración, se le ofrecen sugerencias graduadas y retroalimentación al sujeto sobre los errores cometidos e, igualmente, se le ofrecen estrategias metacognitivas que le ayuden a resolver las tareas propuestas. Se trata en definitiva de «provocar» un proceso de aprendizaje en el alumno evaluado con el fin de obtener información sobre su capacidad para aprender (Guthke y Beckmann, 2000), esperando que esta información proporcione una predicción más adecuada del verdadero potencial intelectual del sujeto.

Los principios fundamentales de este modelo podrían hallarse primeramente en la conceptualización teórica de Vygotski acerca de la ZDP como espacio en el cual, a través de la mediación, tenía lugar el aprendizaje que conduciría a un mayor nivel de desarrollo (Vygotski, 1934/1995; 1979/1996). Así mismo, la concreción de estos postulados por parte de Guthke, está basada a su vez en los posteriores desarrollos del concepto de ZDP que realizó Gal'perin (1966) en su teoría de *la formación sistemática de conceptos y acciones mentales*. En esta teoría, la ZDP es considerada también el lugar idóneo donde se produce el aprendizaje a través de la mediación del adulto. La enseñanza es así contemplada como una actuación en esta «zona» que debe llevar al alumno hacia cotas de desarrollo superiores (Sternberg y Grigorenko, 2003). Gal'perin consideraba que la enseñanza podía ser observada en términos de *experimento psicológico*, donde la variable independiente —el tratamiento o intervención— debía influir sistemáticamente en el desarrollo cognitivo del niño, es decir, en el resultado del aprendizaje, que haría las veces de la variable dependiente. De esta forma, las funciones cognitivas desarrolladas inicialmente en la ZDP con la ayuda ajustada del adulto, son

interiorizadas y asimiladas más tarde por el alumno. El desarrollo de la teoría de Gal'perin se plasmó en un programa educativo general, impregnado en su totalidad de la noción de zona de desarrollo potencial.

Los trabajos desarrollados por Guthke y sus colaboradores (Guthke, 1992; 1993; Guthke y Stein, 1996; Guthke y Wiedl, 1996; Guthke y Beckmann, 2000) han dado lugar a un conjunto de pruebas del potencial de aprendizaje encuadradas metodológicamente en un modelo clásico de pretest/ instrucción /postest y que se ajustan a la distinta duración establecida para la fase de instrucción. De esta forma, se distinguen dos tipos de pruebas: unas que se caracterizan por contar con una mediación más extensa y otras que se desarrollan en un menor espacio de tiempo, intercalando entre las fases inicial y final una instrucción más intensiva y corta. Un tipo de prueba que se caracteriza por contar con una fase de intervención prolongada (7 días), es la *Prueba de Razonamiento*; se trata de una batería que consta de dos pruebas paralelas que pueden ser alternadas en pretest y postest y ha sido diseñada para evaluar el razonamiento verbal, numérico y espacial. Los alumnos evaluados reciben durante la fase de intervención, que está normalizada, una instrucción específica en el uso de las estrategias metacognitivas que le ayudarán a resolver los ítems de la prueba, a través de unos manuales de enseñanza programada. La batería puede ser aplicada individualmente o en grupo, constituyendo el resultado obtenido en el postest la medida del potencial de aprendizaje. Otro tipo de pruebas diseñadas por estos autores suponen, como adelantábamos anteriormente, una fase de instrucción más breve, incluyéndose ésta de una forma más clara en el procedimiento de evaluación propuesto. Estas baterías breves se han desarrollado en dos direcciones: la primera de éstas ofrece retroalimentación sistemática al alumno, pero de forma limitada; el segundo formato que han tomado estas pruebas se caracteriza por ofrecer retroalimentación y ayuda abundante al alumno durante la evaluación mediante unas pautas normalizadas. Según Sternberg y Grigorenko (2003), estas pruebas no contendrían una fase de intervención propiamente dicha, sino que, más bien se trataría de la introducción de variaciones o manipulaciones de la situación de prueba para obtener la información que se desea. Se han desarrollado varias pruebas breves de aprendizaje, algunas de las cuales están basadas en tests convencionales, como la *Prueba de Aprendizaje de Raven* (Guthke, 1992), que fue diseñada para identificar y evaluar precozmente problemas del desarrollo; otras pruebas, en cambio, introducen nuevos ítems elaborados por Guthke y sus colaboradores.

Los principios generales que rigen las pruebas elaboradas por Guthke y sus colegas con un formato de aplicación breve para la evaluación de la capacidad de aprendizaje son los siguientes (Guthke y Beckmann, 2000):

- Garantizar la validez de contenido sobre la base de un análisis cognitivo-psicológico del conjunto de ítems. Esto implica anticipar y determinar el nivel de complejidad objetivo de los ítems en contraste con un análisis subjetivo presente en las teorías clásicas de evaluación.
- Construir secuenciadamente los requisitos de la prueba en función de un análisis de las tareas.
- Introducir sistemáticamente retroalimentación o sugerencias como parte integral del procedimiento de la prueba.
- Proporcionar tareas más complejas después de una resolución independiente o después de haber recibido asistencia del programa para poder solucionar los ítems.

- El concepto Prueba de aprendizaje está conectado con el principio de *enseñanza adaptativa*, en el sentido de proporcionar un proceso individualizado en el que los ítems son presentados en función de las características previas del sujeto y de su actuación durante la prueba.
- Se realiza un análisis cualitativo de los errores cometidos y del tipo de ayuda que produce respuestas más adecuadas.
- Tener en cuenta los procesos mencionados para realizar un diagnóstico del proceso de aprendizaje.

Estas pruebas suponen un intento de conjugar las ventajas y las fortalezas de una aproximación psicométrica estandarizada en la evaluación de la inteligencia, es decir: la posibilidad de ser aplicadas sin una formación prolongada y exhaustiva, la disponibilidad de puntuaciones estándar que sirvan de referencia a los aplicadores o el énfasis en la fiabilidad y la validez de los ítems que componen las pruebas; con las ventajas y fortalezas de aplicar dinámicamente las tareas, con el objetivo de obtener información adicional sobre el potencial de aprendizaje del sujeto. Debido a esto, algunos autores han destacado que los procedimientos del modelo propuesto por Guthke se acercarán más a una tradición psicométrica que a la evaluación dinámica (Haywood, 1997).

Los materiales usados en las Pruebas de aprendizaje se componen, como apuntábamos al comienzo, de algunas pruebas convencionales como por ejemplo las matrices progresivas de Raven, seriación de números y figuras, y analogías verbales. También se han desarrollado nuevos ítems para la variante más breve de estas pruebas. Algunos de estos ítems tampoco guardan relación con el currículum escolar, es decir, se corresponden con tareas habituales en los tests de inteligencia; otros, en cambio, han sido contruidos con base en el currículum y son bastante similares a las tareas y pruebas escolares. Estos ítems fueron desarrollados tras un estudio específico de las habilidades implicadas en la lectura y el cálculo (Guthke y Beckmann, 2000). Las pruebas pueden ser realizadas como tareas de lápiz y papel o mediante un programa informático.

Como decíamos con anterioridad, existen varias pruebas desarrolladas por este grupo de investigadores; entre ellas podemos destacar el LLT (*Leipzig Learning Test*) y el ACIL (*Adaptative Computer Assisted Intelligence Learning Test Battery*) (Guthke, Beckman, Stein, Vahle y Rittner, 1995). El LLT ha sido desarrollado para valorar las habilidades cognitivas y la capacidad de aprendizaje de niños de primer curso de educación primaria y muy especialmente de aquellos que presentan NEE relacionadas con problemas del lenguaje. Esta prueba se considera especialmente útil para la toma de decisiones sobre la modalidad de escolarización idónea para niños con NEE, decisiones sobre la repitencia del alumnado, necesidad de apoyo o tutorización, admisión en Centros especiales para la atención de las dificultades en el lenguaje, etc. El objeto primordial de valoración en esta prueba lo constituyen el razonamiento abstracto y la formación de conceptos (Vygotski, 1934/1995; Luria, 1995). En este sentido, el LLT incluye un sistema de ayudas y sugerencias que van más allá de una simple retroalimentación del tipo acierto/error, como la que está presente en el Test de Clasificación de Tarjetas de Wisconsin (Grant y Berg, 1948), encajando plenamente en una concepción de Prueba de aprendizaje y permitiendo una mayor discriminación de la ayuda requerida por niños y niñas con bajo funcionamiento cognitivo para poder solucionar los problemas. Básicamente la prueba consta de una sección introductoria y otra de actividad con las

tareas. En la primera sección el niño es familiarizado con los objetos que van a ser usados durante la prueba y con sus propiedades. El evaluador recoge información sobre la respuesta espontánea del niño y corrige los errores; se trata, de esta forma, de lograr una similar posición de partida al inicio de la prueba. El LLT consta a su vez de tres subtests: en el subtest I se procede con la clasificación en función de la forma; en el subtest II, en función del tamaño, y en el subtest III se realiza en función de la combinación forma-tamaño. Cada subtest contiene 5 tareas, conformando las tres primeras la fase de entrenamiento y sirviendo como control, las dos finales, en las que el criterio de clasificación aprendido en las tres primeras tareas y el número de objetos trabajado permanece inalterado, cambiando sólo características no relevantes del dibujo.

El ACIL es una batería de Pruebas de aprendizaje diseñada especialmente para su aplicación informática. Consta de tres pruebas de aprendizaje o subtests de duración breve: el ADAFI (*Adaptative Sequential Figure Learning Test*), que es una prueba de secuencias figurales en las que el alumno debe continuar series de complejidad creciente; el AZAFO (*Adaptative Number Sequence Learning Test*), que evalúa la capacidad de continuar distintas series numéricas, también con niveles de dificultad que se van incrementando; y el ADANA (*Adaptative Analogy Learning Test*), que es una prueba de analogías verbales. La prueba ACIL está pensada especialmente para una población de entre 12 y 16 años y es particularmente apropiada según sus autores en casos de (a) adolescentes con una historia de fracaso escolar relacionado con desventaja sociocultural o dificultades de tipo socioemocional, (b) niños y adolescentes con evidentes problemas de ansiedad u otro tipo de trastornos emocionales que auguren resultados negativos en pruebas convencionales de aptitud, (c) alumnos con DA que fracasan en la escuela o alumnado con deficiencia intelectual, (d) toma de decisiones relativas a la vocación o a la modalidad de escolarización idónea, y (e) daño cerebral y cualquier trastorno que implique déficits cognitivo-conductuales como la impulsividad.

En el subtest ADAFI, la secuencia de las 12 figuras que componen la situación de prueba, está basada en una variación regular relacionada con las características de cada figura clasificada en función del color, la forma y la «gestalt». El número de dimensiones que varían determina la complejidad de los ítems, existiendo tres niveles de dificultad. El AZAFO consiste en 24 series numéricas con 7 elementos en cada una. El alumno debe resolver el 8º elemento. La relación entre los números determina en este caso los cuatro niveles de dificultad existentes. Por último, el subtest ADANA para las analogías verbales, cuyo formato de tarea es similar al de otros tests que evalúan esta misma capacidad (Jensen, 2000), representándose habitualmente como: $A:B = C:?$, con cinco alternativas de respuesta para el cuarto término de la analogía. Se presentan al alumno un total de 32 analogías, que contienen relaciones de contraste entre conceptos, comparación, subordinación, coordinación, localización, objeto, de tipo instrumental y de finalidad.

El ACIL es una prueba diseñada especialmente para su uso por parte de psicólogos educativos y profesorado de apoyo. En su aplicación, toda la información posible es recogida, desde las reacciones y respuestas espontáneas del sujeto hasta el tiempo empleado en la actividad; se trata de hacer una reconstrucción lo más fiable posible de la actividad del sujeto con el fin de analizar cuantitativamente y cualitativamente la prueba. De forma concreta, las medidas obtenidas se basan en: (a) el número de tareas completadas, (b) el número de sugerencias aportadas en la aplicación, (c) el número de pasos (entendido como la suma del número de tareas completadas y el número de

sugerencias requeridas, (d) latencia entre la presentación de la tarea y la respuesta del sujeto, y (e) el tiempo empleado en la resolución total de la prueba.

En las nuevas variantes de las Pruebas de aprendizaje, el concepto de aprendizaje está conectado con el principio de *enseñanza adaptativa* (Kubinger, 1986; Coll y Mirás, 2001; Coll, Martín y Onrubia, 2001). Por ejemplo, la prueba ACIL contempla la posibilidad de que los alumnos se «desplacen» por las actividades del programa sin recibir ayuda en un principio hasta que cometan algún error en la resolución de alguno de los ítems. En ese caso, el programa introduce un ítem más fácil, que es propuesto para su resolución, y si comete dos errores seguidos se le ofrecen sugerencias relacionadas con el tipo de error cometido con base en un análisis cualitativo de los mismos; posteriormente, se le ofrecen ítems de complejidad equivalente y si los resuelve satisfactoriamente, se le vuelven a presentar ítems con un nivel de complejidad similar al que causó el error inicial. Se produce de esta forma una continua adaptación de la prueba al nivel de actuación del sujeto. Éste se va enfrentando a tareas fáciles y difíciles de forma «circular» (Sternberg y Grigorenko, 2003) hasta que, o bien alcanza el criterio, resolviendo los ítems propuestos, o bien falla de forma continuada en los ítems presentados, sin realizar ningún avance, o consume el tiempo de que dispone. La dificultad de los ítems es determinada para cada sujeto en función de su edad. La información proporcionada por esta prueba se relaciona con los niveles de dificultad de los ítems inicial y final, es decir, con qué nivel de complejidad comenzó y hasta dónde ha llegado; se relaciona igualmente con el número y tipo de errores cometidos; así como con el número de sugerencias requeridas para solucionar los ítems. Esta prueba permite la individualización del proceso de evaluación sin menoscabar, según sus autores, las posibilidades de estandarización y, por tanto, de comparabilidad de los resultados de la evaluación.

Investigación y revisión crítica del modelo

Las investigaciones que se han llevado a cabo por Guthke y sus colaboradores utilizando las Pruebas de aprendizaje, han tenido por objeto, aparte de establecer las propiedades psicométricas de las distintas pruebas diseñadas, determinar el peso de la predicción de las medidas obtenidas en el postest, con relación al rendimiento académico y en comparación con las puntuaciones pretest; también se han dirigido a comparar la mejora experimentada por los sujetos cuando participaban de la evaluación con Pruebas de aprendizaje —incluso cuando eran utilizadas las versiones con una fase de intervención breve— o cuando lo hacían de una valoración que constaba de la simple repetición del pretest empleado. En general, la investigación parece confirmar que la mejora obtenida en el postest no puede ser predicha totalmente por las puntuaciones pretest, por lo que los autores afirman que la ED con las Pruebas de aprendizaje proporcionaría una información adicional muy relevante para conocer la capacidad de aprendizaje de los sujetos.

En un estudio de Guthke y Wingenfeld (1992) se investigó la validez de predicción que mostraban las puntuaciones postest en una Prueba de aprendizaje, sobre las calificaciones escolares y la valoración del profesorado, en comparación con las obtenidas en el pretest, siendo significativamente más altas las correlaciones obtenidas con las medidas postinstrucción. En este mismo estudio se comparó también la mejora que obtenía un grupo de alumnos experimental al que se había administrado una Prueba de aprendizaje de formato breve en comparación con otro grupo de alumnos control, al

que se le había aplicado el test de matrices progresivas de Raven dos veces; la mejora obtenida en el postest resultó ser significativamente más alta en el GE. Guthke y Wingenfeld también demostraron que el poder predictivo de las Pruebas de aprendizaje era significativamente mayor para el grupo de niños que presentaba un nivel intelectual bajo que para los de nivel medio y superior; de este modo, la información aportada por estas pruebas se tornaba más relevante y decisiva cuando hacía referencia al alumnado con DA, NEE, o que presentaban desventaja sociocultural.

En una investigación realizada en Holanda (Heesels y Hamers, 1993) con grupos de niños inmigrantes y holandeses (N=500), se analizaron las diferencias en las puntuaciones obtenidas en pruebas estáticas y dinámicas. Los resultados mostraron que los niños inmigrantes puntuaban significativamente más bajo que sus compañeros holandeses en las pruebas estáticas; sin embargo, esta diferencia desaparecía cuando la aplicación había sido efectuada con una Prueba de aprendizaje. Los autores del estudio descubrieron que la estructura factorial de las puntuaciones del potencial de aprendizaje era la misma para ambos grupos.

La reciente revisión de Sternberg y Grigorenko (2003) aporta una serie de reflexiones acerca de cuestiones metodológicas relacionadas con los estudios presentados; a estas reflexiones uniremos también algunas nuestras acerca del modelo presentado. En primer lugar, debemos decir que, aunque se ha desarrollado un estudio longitudinal (Guthke, 1992) sobre el valor predictivo de las medidas del potencial de aprendizaje en niños que fueron evaluados en 1º, 2º, 6º y 7º curso, y se habían comparado los resultados obtenidos con la predicción efectuada a través de las medidas obtenidas con la medición estática del mismo test empleado (matrices progresivas de Raven), siendo significativas las diferencias a favor de la aplicación dinámica del test; factores como el tamaño de la muestra (N=28) y la escasa potencia de las correlaciones obtenidas, puede indicar una cierta debilidad de la magnitud de los efectos producidos. Cuando se han realizado estudios con muestras mayores, como la que exponíamos anteriormente realizado en Holanda, los resultados no han mostrado mayor poder de predicción de la Prueba de aprendizaje sobre las pruebas estáticas. Estas contradicciones en los resultados obtenidos —empleando distintas Pruebas de aprendizaje, aunque el mismo procedimiento— parecen indicar la necesidad de otros estudios que, o bien confirmen la validez predictiva de las pruebas, o bien estén dirigidos a incrementar esta propiedad.

Por otra parte, algunas investigaciones del grupo de Guthke se han dirigido a establecer las diferencias entre la información obtenida con tests de inteligencia de carácter estático y las pruebas de aprendizaje. Se pretendía así conocer en qué medida ambos tipos de pruebas podían predecir la adquisición de conocimiento, por una parte, y el producto de esta adquisición de conocimiento, por otra. En un estudio realizado con alumnado de 8º grado (Beckmann y Guthke, 1999), en el que se utilizaron idénticos ítems pero distintos procedimientos, se evaluaron dos tipos de problemas: unos que no guardaban relación con el currículum —inferir el funcionamiento de una máquina— y otros relacionados con las áreas académicas de Geografía y Matemáticas. Las medidas criterio fueron evaluadas un año después de la administración de la prueba. Los resultados mostraron que se producía una interacción entre las varianzas proporcionadas por los tests de inteligencia y por las pruebas de aprendizaje, interacción que explicaba la mayor parte de la varianza. Por otra parte, las pruebas de aprendizaje explicaban una parte de varianza específica del criterio, mientras que las pruebas estáticas no podían hacerlo. Las calificaciones escolares en Geografía y Matemáticas no predijeron

adecuadamente el éxito en el aprendizaje de estas materias como así lo hicieron las pruebas de aprendizaje y las pruebas estáticas de inteligencia. La importante conclusión a la que llegan los autores del estudio hace referencia a que la información que proporcionan los tests de inteligencia estaba incluida en la que proporcionaban las Pruebas de aprendizaje, pero que, por el contrario, la información que podían proporcionar las Pruebas de aprendizaje no se encontraba recogida en los tests de inteligencia (Guthke y Beckman, 2000).

En un estudio presentado por Guthke y Löffler (1983) se utilizó la versión dinámica del Test de Clasificación de Tarjetas de Wisconsin, al que hacíamos referencia anteriormente, con distintos grupos de alumnos de primer curso de primaria y último curso de educación infantil, entre los que había niños con un nivel intelectual dentro de la normalidad, otros que presentaban dificultades de aprendizaje y también otros que habían sido diagnosticados con una disfunción cerebral mínima (DCM) y eran atendidos en el aula de educación especial. En el estudio desarrollado, la dificultad de las tareas fue progresivamente introducida en los ítems presentados. Esta complejidad creciente se manifestaba en la siguiente secuencia: (1) clasificación de 1 elemento en función de la forma (triángulo), (2) clasificación de 1 elemento en función del tamaño (figuras grandes), (3) establecimiento de relaciones entre dos elementos (círculo y figura grande), y (4) clasificación disyuntiva de 2 elementos (triángulo o figuras grandes). Cada uno de los ítems constaba de dos partes: primero era el evaluador quien mostraba la solución, es decir, la clasificación de una serie de conceptos; posteriormente, se pedía al niño que identificara la «regla» o las características que fueron relevantes para la clasificación. Si la respuesta no era correcta, el chico recibía hasta cinco posibles ayudas en función del nivel de complejidad de la tarea. El niño era conducido de esta forma hasta la solución de cada uno de los ítems, mostrándole en última instancia la solución correcta si esto era necesario. Al final de este proceso se pedía al niño que verbalizara las reglas de clasificación utilizadas para resolver el problema. La medida obtenida en la prueba tenía como bases la cantidad de sugerencias requeridas, el tiempo empleado en la resolución y la verbalización durante la actuación.

Uno de los hallazgos, a nuestro juicio, más interesantes del estudio anterior es que a medida que se incrementaba la dificultad de los ítems, disminuía la demanda de ayuda por parte de los niños. El análisis que hacen los autores de este hecho, que podría ser especialmente llamativo a priori en los niños con algún tipo de lesión o disfunción cerebral participantes en el estudio, hace referencia a que éstos, aunque al comienzo mostraban mucha impulsividad y pedían ayuda constantemente, con más asiduidad que el resto de los subgrupos de la muestra; sin embargo, hacia el final de la actividad, donde se encuentran precisamente los ítems más difíciles, estos niños habían comenzado ya a actuar de manera más reflexiva, obteniendo finalmente los mismos resultados que el grupo de niños que no presentaba ningún tipo de dificultad. Este hecho, confirmado por un estudio posterior de Michalski (1988), hizo concluir a los investigadores que la prueba mostraba claramente un potencial de aprendizaje que, sin embargo, no se desvelaba en la actuación escolar de este grupo de chicos. Fue este último autor, Michalski, quien desarrolló una nueva versión de la prueba —que coincide con la actual— que elimina la necesidad de que los niños deban dibujar durante el test, cambiando esta modalidad de respuesta por la utilización de puzzles.

En varios de los estudios realizados se trató de determinar la relación entre el potencial de aprendizaje y otras variables psicológicas. Los resultados que aportan los autores

(Guthke, 1993; Guthke y Wingenfeld, 1992) destacan tres importantes conclusiones: que el potencial de aprendizaje parece ser relativamente insensible a las manipulaciones del entorno, incluido el apoyo recibido por parte de los padres; que tiende a relacionarse con un pensamiento creativo; y que influye en la reducción de la influencia de aspectos socioemocionales durante el desarrollo de las pruebas. A tenor de estos resultados, especialmente del punto que hace referencia a la posibilidad de una cierta insensibilidad del potencial de aprendizaje con respecto al entorno, lo que podría ser en cierto modo contradictorio con otras formulaciones del potencial de aprendizaje especialmente sensibles a la manipulación y modificación del entorno, sobre todo en lo referente a la mediación que pueden ejercer las personas de referencia (como los padres) en el aprendizaje de la experiencia (Feuerstein, 1991, 1998a); parece ser que en la medida en que el dispositivo de evaluación utilizado se aleje más de una valoración cualitativa del proceso de aprendizaje y se centre en una obtención de medidas cuantificadas en las que, lógica e inevitablemente, podemos perder cierta información, se haría más difícil valorar la posible influencia del entorno inmediato en los niveles del potencial de aprendizaje. Por otra parte, la metodología utilizada para establecer estas asociaciones parece exclusivamente correlacional, lo que nos llevaría a pensar sobre qué medida cuantitativa se ha empleado para determinar la influencia del entorno y si ésta era idónea o suficiente.

Para concluir, destacar el interés primordial de Guthke y sus colaboradores por evaluar el proceso y la capacidad de aprendizaje de los sujetos, ideando para ello diversas situaciones de prueba con fases de intervención de distinta duración y que recogen a su vez contenidos procedentes, tanto de pruebas convencionales de inteligencia, como también otros más relacionados con el contexto escolar. Igualmente debemos destacar el énfasis puesto desde este modelo en la normalización del proceso de evaluación, tratando de incrementar la fiabilidad de las puntuaciones obtenidas. Este proceso de normalización y estructuración de la fase de intervención se ha tratado de conjugar con los principios de adaptación y flexibilidad en la actividad realizada que caracterizan la enseñanza adaptativa —así como también la ED— y ha tenido su desarrollo más relevante a nuestro parecer en la elaboración y aplicación experimental de la Prueba ACIL, que hemos tenido ocasión de conocer. Como ya apuntábamos anteriormente, el análisis cualitativo del proceso de aprendizaje durante su aplicación, puede ofrecer información adicional muy valiosa a la ya recogida en la prueba mediante la valoración cuantitativa, incrementando de esta forma también *cualitativamente* la información que es posible obtener con respecto a la aplicación de pruebas convencionales.

La Prueba de Procesamiento cognitivo de H.L. Swanson (TPC-S)

Bases teóricas y metodológicas

Esta prueba dinámica desarrollada por Swanson (1995a, 1995b) halla sus bases teóricas en la psicología del *procesamiento cognitivo de la información*. Este autor se propone ofrecer un instrumento capaz de responder a la necesidad de evaluar componentes significativos de las habilidades de procesamiento mediante un modelo dinámico de evaluación (Swanson, 2000). Al igual que otros investigadores de las pruebas de ED de las que se hace una revisión en este trabajo, una de sus máximas preocupaciones es dotar a este modelo de propiedades psicométricas mediante la estandarización del proceso de evaluación; otra preocupación la constituye el hecho de elaborar un instrumento que pueda ser útil y eficaz para el trabajo que deben desarrollar

los psicólogos educativos en los centros escolares. El componente de la capacidad de procesamiento que se evalúa con el TPC-S es la memoria de trabajo (MT); se trata pues de uno de los instrumentos de ED que dirige su atención de una forma más clara hacia procesos concretos implicados en el aprendizaje y no hacia el potencial de aprendizaje en conjunto. El autor basa la decisión de focalizar la evaluación sobre la MT en que este constructo ha sido considerado como un proceso esencial para el aprendizaje y ha sido incluido en la mayoría de los modelos de enseñanza y aprendizaje de estrategias elaborados desde la teoría del procesamiento de la información (Bruning, Schraw y Ronning, 2002; Baddeley, 1996; Just y Carpenter, 1992), correlacionando de forma significativa con la mayoría de las actividades académicas relacionadas con el lenguaje y las matemáticas. La MT se relaciona de esta forma con las DA, de forma que éstas son atribuidas a déficits en la capacidad de mantener y operar simultáneamente información requerida para desarrollar una actividad o solucionar un determinado problema (Swanson, 1995c).

El modelo de Swanson relaciona la MT y la memoria a largo plazo (MLP) en la explicación del modo en que se produce la codificación de la información durante la prueba; de esta forma, cuando una información es codificada en la MT y, por tanto, es posible operar con ella, ha debido ser previamente codificada en la MLP; dicho de otro modo, cuando las representaciones de la MLP intervienen en el proceso de la actividad como resultado de aprendizajes anteriores, se produce la codificación en la MT (Swanson, 2000).

El TPC-S pretende ofrecer un indicador aproximado del *procesamiento potencial*, entendido éste como la diferencia entre un nivel de procesamiento actual, medido individualmente como se hace en un test convencional, y el nivel de competencia en el procesamiento de la información que puede lograr el niño con la ayuda del evaluador. La capacidad de incorporación o apropiación de las sugerencias y ayudas ofrecidas a la propia actuación con el fin de mejorarla, determina el nivel de procesamiento potencial. Esta apropiación del sujeto con base en la mediación del evaluador proporcionaría una valiosa información sobre si el acceso a la información tiene lugar, o si éste puede ser facilitado con vistas al logro de un mayor nivel de procesamiento. Desde este ángulo de visión sobre la ED, la información obtenida tras el proceso de evaluación consistiría esencialmente en la valoración de la capacidad del sujeto para mantener en su MT un nivel de actuación superior, como el proporcionado durante la fase de intervención.

Por otra parte, el TPC-S se compone de una serie de subpruebas que implican el mantenimiento y el procesamiento simultáneo de la información. Estas subpruebas, once en total, son aplicadas mediante el uso de distintas modalidades (cuadro 19); de este modo, hay actividades con un formato más visual, mientras que otras presentan una modalidad auditiva; también existen tareas con contenido verbal y otras con contenidos no verbales; por último, algunas actividades evalúan aspectos de forma retrospectiva, mientras que otras lo hacen de forma prospectiva. Esta diversidad de tareas y de modalidades de intervención y respuesta, pretende adecuarse a las exigencias de evaluar dinámicamente un aspecto como la MT, que se considera central en el procesamiento y codificación de la información de cualquier contenido.

El contenido de estas subpruebas está desarrollado en su totalidad para la valoración de la MT, pero puede a su vez ser clasificado en función de varios criterios (Swanson, 2000). En primer lugar, se distingue entre memoria *episódica* y *memoria semántica*.

Cuadro 19. Descripción de los subtests que componen el TPC-S. Tomado de Swanson, 2000

<i>Subtest 1</i> Rimas	Consiste en la evaluación del recuerdo de grupos de palabras que riman de forma progresivamente más compleja.
<i>Subtest 2</i> Matriz visual	Evalúa la capacidad de recuerdo de secuencias visuales dentro de una matriz.
<i>Subtest 3</i> Serie auditiva de números	Evalúa la capacidad de recuerdo de la información numérica contenida en una pequeña frase.
<i>Subtest 4</i> Mapas y direcciones	Evalúa la capacidad para recordar una serie de direcciones sobre un mapa que no tiene nombres.
<i>Subtest 5</i> Repetición de relatos	Evalúa la capacidad para recordar una serie de episodios relatados en un párrafo.
<i>Subtest 6</i> Serie de imágenes	Evalúa la capacidad para recordar una serie de formas progresivamente más complejas espacialmente.
<i>Subtest 7</i> Serie de frases	Evalúa la capacidad de recuerdo de frases aisladas como p.e.: el río pasa; el oso lento; el niño crece...
<i>Subtest 8</i> Organización espacial	Evalúa la capacidad para recordar la organización espacial de tarjetas que contienen imágenes con varias formas ordenadas de mayor a menor.
<i>Subtest 9</i> Asociación semántica	Evalúa la capacidad para organizar palabras dentro de categorías abstractas, debiendo tomar aquellas que van juntas, p.e.: camisa, pantalón y zapatos; sierra, martillo y puntillas; etc.
<i>Subtest 10</i> Categorización semántica	Evalúa la capacidad de recordar asociaciones de palabras dentro de unas categorías, p.e.: profesiones: maestro, bombero, policía; estaciones: verano, invierno, otoño; etc.
<i>Subtest 11</i> Secuenciación no verbal	Evalúa la capacidad para secuenciar una serie de tarjetas con imágenes de formas no conocidas cuya organización no es proporcionada por el evaluador.

Un criterio utilizado para valorar estos aspectos en los subtests está en función de la naturaleza de los elementos recordados, el TPC-S distingue entre varios subtests que se centran en el recuerdo y manipulación de elementos aislados, lo que haría referencia a la memoria episódica, y otros que se centran en el recuerdo y manipulación de categorías y/o clasificaciones, lo que nos ofrecería información sobre la memoria semántica. Otro criterio sería el que hace referencia a la naturaleza de la relación entre los elementos; en este sentido, la memoria episódica se fundamentaría en las secuencias de elementos o acciones, y la semántica en la organización de los elementos. En segundo lugar, los subtests y sus ítems pueden ser clasificados en función de los errores o fallos de memoria cometidos. Se distingue entre *fallos de memoria retrospectiva* y *fallos de memoria prospectiva*; en los primeros, el sujeto olvida incluir elementos relevantes en su respuesta, lo que podría estar causado por (a) una falta de control consciente del proceso, control que sería requerido para poder dar solución a la tarea, o también por (b) focalizar la atención en información no relevante o diferente a la que se pide. Los subtests 1, 2, 5, 6 y 9 contienen ítems que pueden dar lugar a los fallos mencionados. Por otra parte, los fallos de memoria prospectiva se cometerían cuando el alumno no selecciona y aplica adecuadamente las estrategias necesarias para categorizar y organizar la información, estrategias que le serán necesarias para recordar. Los subtests 3, 4, 7, 8, 10 y 11 contienen ítems que evalúan estos posibles fallos. En tercer lugar, otra

posible distinción es la que hace referencia a la modalidad de procesamiento requerido: de esta forma, se distingue entre tareas que exigen un procesamiento verbal (escucha y evocación verbal: subtests 1, 3, 5, 7, 9 y 10), y tareas que requieren un procesamiento de tipo visual-espacial (manipulación de imágenes y formas: subtests 2, 4, 6, 8 y 11). De todos estos elementos en los que podemos descomponer las actividades propuestas: memoria episódica y semántica, memoria retrospectiva y prospectiva, procesamiento verbal y viso-espacial, se obtienen índices cuantitativos a través del TPC-S.

En lo que respecta al procedimiento de mediación, el TPC-S se basa en proporcionar sugerencias o ayudas que están normalizadas para cada uno de los ítems de la prueba y que son ofrecidas al alumno en función del error cometido. Los once subtests del TPC-S pueden ser administrados como una batería o también por separado. La administración completa puede llevar unas tres horas (Sternberg y Grigorenko, 2003). La medida del *potencial de procesamiento* se compone de hasta 7 puntuaciones: (a) *puntuación inicial*, interpretada como el nivel de procesamiento actual del sujeto y que viene a ser equivalente a las medidas obtenidas con tests estáticos, (b) *puntuación de mejora*, que sería la puntuación máxima obtenida por el sujeto durante las pruebas, (c) *puntuación de eficacia instructora*, que hace referencia al número de sugerencias necesarias para lograr el nivel máximo de procesamiento, (d) *puntuación de mantenimiento*, que es interpretada con base en la estabilidad del nivel de procesamiento obtenido sin el apoyo del evaluador, (e) *puntuación de diferencia de procesamiento*, que hace referencia a la valoración de la diferencia entre el nivel de procesamiento inicial y el nivel de procesamiento potencial, determinado en condiciones de colaboración y ayuda del evaluador, (f) *puntuación de estabilidad de procesamiento*, que se determina por la diferencia entre la puntuación inicial y la puntuación de mantenimiento, y (g) *puntuación de eficiencia estratégica*, que hace referencia a las estrategias de memoria que deben seleccionar y usar los alumnos evaluados para poder realizar algunos ítems. Esta última puntuación trataría de reflejar un conocimiento declarativo de estrategias de memoria.

El TPC-S pretende ofrecer información útil al profesorado de cara a la mejora del nivel de procesamiento de la MT. Esta información es relativa a los procesos de intervención considerados idóneos en función de los resultados de la aplicación de la prueba. El enfoque de la información proporcionada a los educadores se realiza en torno a cuatro parámetros: (1) efectividad de una retroalimentación simple sobre la actuación del alumno, (2) conocimiento procedimental, (3) el grado en que el nivel alcanzado por el alumno se mantiene una vez retirada la ayuda del evaluador, y (4) modalidad preferente para presentar la información (verbal o visual-espacial), a través de la cual el profesorado puede conocer el tipo de ayuda necesitada con el objetivo de ajustar ésta y optimizar el mantenimiento de la información.

El instrumento desarrollado por Swanson fue concebido en un contexto educativo de atención a las NEE que presentaban alumnos y alumnas con problemas en el aprendizaje de la lectura y las matemáticas principalmente. La hipótesis de que la MT subyacía a estas dificultades de aprendizaje hizo que en los estudios planteados desde este modelo se abordaran cuestiones relacionadas con el posible uso diferencial que de estos procesos podían hacer estos niños en comparación con aquellos que no presentaban problemas, así como con la posible especificidad del grupo de niños que presentaban estos déficits en su MT con respecto a otros que también presentarían dificultades en su aprendizaje, pero no mostrarían especiales problemas en su MT. Por

otra parte, en este contexto se pretendía estudiar el grado de modificabilidad de la MT para estos niños con dificultades mediante la aplicación de la prueba dinámica desarrollada.

Investigación y revisión crítica del modelo

Del proceso de estandarización del TPC-S, llevado a cabo por Swanson y sus colaboradores en Estados Unidos y Canadá durante varios años y contando con una muestra de 1611 sujetos, muy diversificada en cuanto a factores como la edad, el nivel sociocultural y económico, la condición o no de discapacidad, el origen cultural y étnico, o el lugar de residencia, urbano o rural; pueden extraerse algunas conclusiones que nos pueden resultar relevantes más allá de la determinación de propiedades psicométricas. En primer lugar, del análisis factorial realizado con base en las *puntuaciones iniciales, puntuaciones de mejora y de mantenimiento*, es decir, las que son obtenidas a modo de pretest y postest (inicial y mantenimiento respectivamente), sin el apoyo de las sugerencias del evaluador, y la que indica la puntuación de la fase de intervención (mejora), se obtienen dos factores que son interpretados por el autor como una medida de la memoria semántica (Rimas, Serie auditiva de números y Asociación semántica) y una medida de la memoria episódica (Matriz visual, Mapas y direcciones y Repetición de relatos). Este análisis es representativo de los datos obtenidos de la pasación completa del test, así como de la forma abreviada, compuesta por las 6 subpruebas mencionadas. En segundo lugar, se efectuaron una serie de análisis correlacionales entre la puntuación compuesta del TPC-S y varias medidas de tests estandarizados convencionales. Estos análisis muestran que las puntuaciones: *inicial, mejora y mantenimiento* de la prueba son más sensibles que las del resto de componentes a las medidas de tests de procesamiento, inteligencia y académicos. Por ejemplo, la puntuación inicial correlacionaba 0,80 con el test K-ABC de Kaufman; de 0,38 a 0,88 con el WISC-R de Wechsler y 0,60 con el test de Raven; sin embargo, las medidas correspondientes a los componentes de *eficacia instructora, estabilidad de procesamiento y diferencia de procesamiento*, mostraban bajas correlaciones con los tests de inteligencia y de logro académico. Los datos anteriores son interpretados por el autor en el sentido esperado de que algunos componentes del potencial de procesamiento pueden estar débilmente correlacionados con medidas tradicionales y estáticas. Por otra parte, se esperaba que el componente de la *eficiencia estratégica*, que hacía referencia al conocimiento declarativo de estrategias, estuviera relacionado con indicadores de metacognición como la conciencia sobre el propio conocimiento y sobre la capacidad de procesamiento; así mismo se esperaba que no lo estuviera con las medidas psicométricas estándar, cómo así ocurrió, por las mismas razones expuestas anteriormente. El conocimiento de estrategias y, más aún, el metaconocimiento estratégico, no es un elemento que esté presente directamente en los tests convencionales, de ahí que, a pesar de que pueden hallarse evidencias de su relevancia para la ejecución práctica, como así ha ocurrido en otros estudios citados en nuestro propio trabajo (Vidal-Abarca y Gilabert, 1991; Alonso Tapia, Carriedo y Mateos, 1992; Montanero, 2001; Navarro y Mora, 2003a), no es extraño para nosotros tampoco el hecho de que las puntuaciones de tests convencionales no correlacionen con puntuaciones de medidas de conocimiento o metaconocimiento estratégico, ya que en la mayoría de los casos, el constructo evaluado por estos tests puede no tener relación alguna con la eficacia del sujeto a la hora de conocer y seleccionar la estrategia más adecuada para resolver una determinada tarea.

En cuanto a las correlaciones con pruebas de rendimiento académico, las *puntuaciones iniciales, de mejora y de mantenimiento* logran correlaciones altas, en torno a 0,53 con el subtest de lectura del WRAT (*Wide Range Achievement Test*) (Jastak y Jastak, 1978; Jastak y Wilkinson, 1984) y en torno a 0,60 con el de logro académico. En los subtests de lectura de reconocimiento, lectura, y comprensión del PIAT (*Peabody Individual Achievement Test*) (Dunn y Dunn, 1981), las correlaciones para las tres puntuaciones componentes del TPC-S mencionadas son también muy altas, en torno a 0,60; igualmente altas, de 0,47 a 0,63, son las correlaciones obtenidas especialmente por estos componentes con los subtests de un test de lenguaje escrito, TOWL-2 (*Test of Written Language-2*) (citado en Swanson, 2000). Con el SDRT (Stanford Diagnostic Reading Test), contando con una muestra compuesta por 30 alumnos con DA de edades comprendidas entre los 9 y los 13 años, los resultados de las correlaciones entre el subtest de comprensión lectora y las distintas puntuaciones componentes del TPC-S son estas: 0,62 para la puntuación inicial, 0,56 para la de mejora, 0,63 para la de mantenimiento, 0,33 para la puntuación de la diferencia de procesamiento, 0,20 para la de eficacia instructora, 0,38 para la de estabilidad de procesamiento y -0,08 para la de eficiencia estratégica. Si bien las correlaciones que mantienen las puntuaciones inicial, mejora y mantenimiento, son altas; por el contrario, las correlaciones obtenidas por el resto de las puntuaciones componentes del TPC-S con los test de inteligencia y de rendimiento académico citados, son marcada y significativamente más bajas. En cambio, esta tendencia se invierte ante el contraste efectuado con un Cuestionario de Metacognición (Swanson y Thahan, 1992; Paris et. al., 1984). Las correlaciones con este cuestionario, para una muestra de 31 alumnos de 8 a 15 años, arrojaron los siguientes resultados: 0,05 para la puntuación inicial, -0,23 para la de mejora, 0,17 para mantenimiento, -0,24 para la de diferencia de procesamiento, -0,29 para la eficacia instructora, -0,28 para la estabilidad de procesamiento y 0,66 para la puntuación de eficiencia estratégica.

Otro interés manifiesto en los estudios de Swanson ha sido determinar hasta qué punto las clasificaciones del alumnado efectuadas con base en su funcionamiento cognitivo son estables. De acuerdo con una concepción dinámica de estas funciones, apoyada en numerosos trabajos (Campione, Brown y Ferrara, 1987; Campione, Brown, Ferrara, Jones y Steinberg, 1985; Haywood, Brown, Wingensfeld, 1990; Feuerstein, 1991; Sternberg y Grigorenko, 2003), la ED puede llegar a modificar significativamente la capacidad manifiesta de un sujeto para enfrentarse a una situación problemática y, por tanto y en consecuencia, también la valoración que de su funcionamiento cognitivo, y especialmente de su capacidad de aprendizaje, podemos hacer. Swanson se propuso determinar la contribución del TPC-S a estos posibles cambios en las clasificaciones basadas en pruebas estáticas. Para ello comparó un grupo de alumnos identificados en la escuela como alumnos que presentaban DA en su lectura con un grupo de alumnos identificados como «lectores estratégicos». La comparación se efectuó mediante la realización de un conjunto de subtests de la prueba TPC-S aplicados de forma estática y dinámica (Swanson, 1995a). La muestra, que no difería significativamente en edad o género, aunque sí lo hacía en procedencia de tipo sociocultural, estaba compuesta por 155 alumnos con DA en lectura y 351 alumnos que no presentaban problemas. La muestra de alumnos con DA respondía a las características y la definición consideradas prototípicas de esta categoría, es decir: dificultades de aprendizaje específicas y generalmente circunscritas a algún área académica; bajo rendimiento que no es equiparable a su capacidad en otras áreas, en las cuales puede llegar a tener un rendimiento normal; dificultades que no pueden ser explicadas por efecto de déficit

cognitivo, inadecuada o insuficiente enseñanza o deprivación sociocultural. En el estudio se trató de identificar distintos subgrupos entre el grupo de alumnos con DA con base en las puntuaciones de mejora y de eficacia instructora (número de sugerencias requeridas para la solución del problema). Se trataba de averiguar en qué medida los alumnos con DA que mostraron significativamente un más bajo rendimiento en la prueba de procesamiento inicial, que hace las veces de pretest, incrementaban este rendimiento beneficiándose de la mediación en forma de asistencia y sugerencias del evaluador; de esta forma, se pretendían establecer posibles diferencias entre este subgrupo de alumnado con DA en la lectura y déficits en su MT, y otros subgrupos identificados con problemas de aprendizaje, pero que no respondían a las características reseñadas más arriba. El procedimiento utilizado analizó las puntuaciones en 4 dimensiones para cada muestra (con y sin DA) e identificó 4 subgrupos en función de sus puntuaciones en cada una de estas dimensiones: lectura, aritmética, puntuaciones de mejora y puntuaciones de respuesta a la intervención (eficacia instructora). Para la muestra de alumnado con DA, el primer subgrupo mostró una discrepancia mínima entre la puntuación del TPC-S y la puntuación en una prueba de rendimiento en lectura; este grupo puntuaba muy bajo en las 4 dimensiones analizadas; su puntuación de mejora estaba al mismo nivel que su puntuación en lectura y matemáticas; este grupo se definió como de «alumnos lentos». El segundo subgrupo no mostró discrepancias significativas entre la actuación en las pruebas de matemáticas y en el TPC-S, pero sí mostró claras discrepancias con relación a la lectura; estas puntuaciones del rendimiento en lectura son características de los alumnos catalogados como «disléxicos», es decir, alumnos con dificultades específicas en lectura, pero que pueden mostrar un desempeño aceptable en otras áreas académicas (Fletcher, Francis, Rourke, Shaywitz y Shaywitz, 1992; Sánchez, 1999; Moreno y García, 2000). En general, su puntuación de procesamiento permaneció relativamente estable con la evaluación estática y dinámica. El tercer subgrupo mostró una gran discrepancia entre la puntuación de rendimiento, la de mejora y la puntuación de mantenimiento cuando fue comparada la puntuación inicial con la de rendimiento. Este subgrupo obtuvo las puntuaciones de mejora y las de eficacia instructora más altas en comparación con los otros subgrupos de la muestra de alumnado con DA. Este tercer subgrupo se definió como alumnos con «déficits de instrucción», debido a la característica de haber incrementado significativamente sus puntuaciones como respuesta positiva a la intervención. Por último, el cuarto subgrupo mostró una gran discrepancia entre la puntuación del TPC-S y la puntuación de rendimiento, tanto para la lectura como para las matemáticas; sin embargo, su puntuación de mejora y su puntuación de eficacia instructora estaban por debajo de la media; el grupo fue operativizado como alumnado con «DA en lectura y matemáticas».

El anterior análisis permitió reclasificar a un 17% de alumnos (los que conformaban el 3º subgrupo) que debían ser considerados alumnos con *déficits de instrucción* más que con DA; otro 9% de alumnos, del 1º subgrupo, fueron también mejor clasificados en la categoría de alumnos que aprenden lentamente, que como alumnos con DA. Este estudio muestra nuevamente las limitaciones de una valoración estática de las funciones cognitivas y del rendimiento académico en áreas curriculares; la información proporcionada podría indicar la necesidad de reorientar la intervención educativa para ciertos subgrupos de alumnos clasificados en la escuela como alumnado con DA; muestra así mismo las posibilidades de modificación del nivel de procesamiento cuando la evaluación contempla una fase de intervención en la que se ofrece apoyo al alumno en la realización de la actividad. En cuanto a los resultados en términos de *estabilidad de procesamiento*, lo que indicaba la diferencia entre el nivel máximo de procesamiento

alcanzado sin ayuda y la puntuación inicial, los resultados muestran que el nivel de estabilidad fue medio para el subgrupo de alumnos lentos, disléxicos y con problemas de lectura y matemáticas, mientras que para el subgrupo de alumnos recategorizado como alumnos con déficits de instrucción, los niveles de estabilidad de procesamiento indican una alta modificabilidad de su potencial de procesamiento. Swanson informa también acerca de que este subgrupo puntuó al mismo nivel que uno de los que conformaron la muestra de alumnado sin DA, el catalogado como «instruccionalmente eficiente», en las puntuaciones de mejora y de respuesta a la intervención, así como en otras medidas no usadas directamente en la formación de los subgrupos, como las puntuaciones de mantenimiento, diferencia de procesamiento y estabilidad de procesamiento; este 8º subgrupo fue el que más alto puntuó precisamente en respuesta a la intervención entre el alumnado sin DA y el que obtuvo la segunda mejor puntuación media en mejora; este hallazgo es sumamente relevante, ya que permitió asegurar que muchos alumnos considerados con DA y clasificados en consecuencia desde la escuela, poseen un procesamiento potencial de la información equiparable al de sus compañeros que no presentan problemas en su aprendizaje.

Una de las fortalezas de esta prueba la constituye sin duda el hecho de evaluar un proceso cognitivo que se encuentra directamente relacionado con el desempeño de actividades escolares. Los estudios ofrecen además información sobre las relaciones entre la ED propuesta y el rendimiento en lectura y matemáticas, integrando de esta forma el estudio del proceso de MT en el análisis de los procesos de mejora del funcionamiento cognitivo de los niños con DA y NEE. Desde un punto de vista congruente con la teoría del procesamiento de la información, la ED de la MT ofrece potencialmente información valiosa para la intervención, ya que nos daría pistas sobre la capacidad de procesamiento del sujeto y sobre su modificabilidad, orientando la intervención en el sentido de la actuación del sujeto en las distintas tareas; sin embargo, este planteamiento no está exento de problemas. Hemos visto durante la revisión que efectuábamos en el capítulo uno, sobre la lectura y las distintas concepciones que de ella se tienen, que entendida ésta como proceso social-cultural dinámico, de interacción y construcción del conocimiento, difícilmente podía ser desintegrado en subhabilidades sin correr el riesgo de perder el conjunto que da sentido al concepto. Probablemente, ya no haya muchos defensores a ultranza del entrenamiento aislado de subhabilidades que componen la habilidad lectora con el objetivo de mejorar ésta, y no parece en absoluto ésta la posición defendida por Swanson; pero creemos necesario tener cierta cautela con la interpretación que podemos hacer de los resultados obtenidos en los estudios con el TPC-S; los niveles deficitarios de procesamiento y la indicación positiva de modificabilidad de estos niveles deberían llevarnos sólo a la conclusión de que existen niveles deficitarios de procesamiento de la MT y que estos pueden ser mejorados con determinado tipo de intervención, pero no a una conclusión errónea de que los problemas de lectura se solucionarán si entrenamos aisladamente la MT; en primer lugar porque nada sabemos de otros procesos implicados en la lectura y la comprensión, que en esta prueba no son evaluados, y, en segundo lugar, porque parece adecuada una intervención más integral que, aún contando con la información privilegiada que otorga la prueba sobre la capacidad de procesamiento del alumno, y utilizando esta información en la implementación de actividades encaminadas a optimizar el nivel de procesamiento simultáneo de la información; tenga en cuenta el texto como unidad de significado y la intervención e interacción de distintos procesos implicados.

Otro de los aspectos especialmente interesantes con relación a nuestro propio estudio es el que tiene que ver con los trabajos llevados a cabo por Swanson para establecer el nivel predictivo del TPC-S sobre el rendimiento en lectura. En un estudio realizado con 61 alumnos (Swanson, 1995a), se utilizó un análisis de regresión por pasos para determinar qué porcentaje de varianza explicada tenían varias puntuaciones: puntuación inicial, puntuación de mejora y puntuación de mantenimiento del TPC-S, y CI. Cuando fueron introducidas en la ecuación las tres puntuaciones compuestas y el CI para predecir el rendimiento en lectura, los resultados indicaron que el 14% de la varianza era atribuida a la puntuación de mejora, el 11% al CI y el 4% a la puntuación de mantenimiento; la puntuación inicial no tenía un peso significativo.

Se han realizado algunas objeciones desde un punto de vista estadístico y metodológico a las aportaciones de Swanson. En concreto, Sternberg y Grigorenko (2003) examinan los procedimientos estadísticos comúnmente empleados para determinar la validez predictiva del TPC-S, así como la idoneidad de la prueba para aportar información adicional que permita clasificar u orientar mejor a los alumnos con DA en la lectura o en las matemáticas. El uso del análisis de regresión por pasos y del análisis de discriminación por pasos para estos efectos parece conllevar cierta inestabilidad de las F obtenidas y cierto sesgo en la estimación de los parámetros estadísticos. Por otra parte, las intercorrelaciones presentadas entre las puntuaciones: inicial, de mejora y de mantenimiento del TPC-S, hacen tambalearse análisis posteriores efectuados para determinar aspectos de validez predictiva. Finalmente, con relación a los estudios sobre la validez de predicción del test, el mismo Swanson introduce otro importante elemento que debe ser investigado aún; éste consiste en la evaluación del nivel predictivo de la prueba con respecto a un proceso de enseñanza prolongado.

En definitiva, a partir de todos los datos que hemos podido revisar y analizar, y a pesar de las objeciones que, tanto otros autores como nosotros mismos hemos realizado sobre diversos aspectos del modelo propuesto por Swanson, debemos destacar, como hemos apuntado anteriormente, el interés que este modelo despierta en nosotros por cuanto nace y se desarrolla en un contexto de atención a la diversidad, preocupado igualmente por ofrecer un instrumento que se muestre eficaz en la detección de posibles dificultades en procesos implicados de forma significativa en el aprendizaje escolar, como es la MT, así como en la mejora de estos mismos procesos mediante la provisión de información al profesorado basada en los datos del proceso de aplicación de la prueba desarrollada. El modelo de Swanson realiza por tanto una contribución muy importante al estudio de la ED; especialmente destacable es el interés mostrado por la distinción de distintos subgrupos dentro del grupo de alumnos con DA en función, principalmente, de su respuesta a la intervención; otro aspecto relevante lo constituye el apoyo que ofrecen estos estudios a la tesis de que las medidas dinámicas aportarían información *adicional* a las efectuadas desde posiciones convencionales y estáticas, y que junto con éstas, ofrecerían una valiosa y completa información del funcionamiento cognitivo del sujeto.

El modelo de Prueba de los Límites de J. Carlson y K.H. Wiedl

Bases teóricas y metodológicas

El modelo de evaluación presentado por Carlson y Wiedl (1980, 1992a, 1992b) está basado en la consideración de tres elementos fundamentales: *el individuo, el procedimiento de evaluación y la tarea*. Estos tres elementos deben ser el eje a través

del cual representar las principales variables a considerar en el proceso de evaluación, así como en la explicación de los cambios en el funcionamiento cognitivo. Esta consideración constituye el modelo teórico en el que los autores basaron la construcción de la Prueba de los límites (*Testing the limits*). Estos investigadores piensan que el contenido y las características de los tests influyen decisivamente en la actuación de los sujetos y que estas características difieren de forma importante entre pruebas que supuestamente miden el mismo constructo, o incluso entre los diversos ítems de una misma prueba. De esta forma, la actividad mental y física que es necesario realizar o que es permitida en los diversos tests, condiciona de manera importante la ejecución del sujeto.

En la descripción de los factores relativos al individuo, desde esta perspectiva se considera la intervención e interacción de tres elementos: (a) *estructuras*, que hace referencia a las estructuras de conocimiento general o específico (epistémicas), y a las estructuras de tipo procedimental (heurísticas); (b) *componentes*, que incluye variables cognitivas y metacognitivas; y (c) *procesos*, que hace referencia a los procedimientos que un sujeto emplea en la resolución de un problema. En cuanto al procedimiento de evaluación o aproximación diagnóstica, se describen tres tipos de enfoques que potencialmente afectarían a la ejecución del sujeto: la modificación, la compensación y la inhibición. La *modificación* incluye cambios a nivel de las estructuras, los componentes y los procesos contemplados en los factores relativos al individuo; la *compensación* no implica cambios de tipo estructural en la ejecución del sujeto, pero supone una intervención que puede optimizar el rendimiento en una tarea determinada mediante la compensación de estrategias o habilidades; por último, la *inhibición* hace referencia al control sobre los aspectos irrelevantes de la actividad. Por lo que respecta a la tarea, en la Prueba de los límites se utilizan principalmente las matrices progresivas de Raven y el test de factor G de Cattell. Estas dos pruebas son la base para descubrir características de la tarea, del individuo o de la situación misma de evaluación que puedan mejorar la ejecución de los sujetos en los tests mencionados. Se enfatiza así mismo la normalización de la intervención del evaluador.

La Prueba de los límites centra su atención en la situación de prueba; la idea fundamental de su conceptualización es que ciertas modificaciones o manipulaciones introducidas en la situación de prueba para compensar dificultades intelectuales o de aprendizaje, pueden incrementar de forma significativa el rendimiento de los sujetos (Carlson y Wiedl, 1979). En este sentido, esta conceptualización de las situaciones de prueba, deudora a su vez de las propuestas iniciales de Schmidt (1971), es también muy similar al diseño de investigación propuesto por Eric Bronfenbrenner (1987); el *experimento ecológico*, como procedimiento para estudiar en entornos naturales eventos o situaciones que pueden tener lugar en raras ocasiones de forma espontánea, consistía básicamente en la introducción de modificaciones en una situación natural para investigar de forma más controlada la respuesta de los sujetos o provocar la aparición de determinadas respuestas que fueran consideradas relevantes por el investigador. Mas adelante volveremos sobre esta idea, cuando abordemos el desarrollo de nuestro estudio. En definitiva, de esta forma, la propuesta de Carlson y Wiedl implica ajustar la manipulación de la situación de prueba al tipo de dificultades de aprendizaje del alumno con el objetivo de proporcionar las mayores posibilidades de obtención de información sobre este aprendizaje.

Cuadro 20. Esquema conceptual para el análisis de la ejecución en la prueba y del cambio.
Tomado de Carlson y Wiedl, 2000

<p>I Características de la tarea</p> <p><i>Niveles</i></p>	<ul style="list-style-type: none"> • Actividad práctica • percepción directa • percepción indirecta • conocimiento conceptual lingüístico
<p>II Factores relativos al individuo</p> <p><i>Estructuras</i></p> <p><i>Componentes</i></p> <p><i>Procesos</i></p>	<ul style="list-style-type: none"> • epistémica • heurística • capacidades básicas (relaciones, seriación, clasificación) • capacidades procedimentales (planificación, flexibilidad, autorregulación, activación...) • variables de orientación (orientación específica de la tarea, autoconcepto, conceptualización de la situación...) • análisis • síntesis
<p>III Aproximación diagnóstica</p> <p><i>Modificación</i></p> <p><i>Compensación</i></p> <p><i>Inhibición</i></p>	<ul style="list-style-type: none"> • implícita/explicita • predeterminada/autodeterminada • protésica (implica apoyo para la resolución de una actividad) • catalizadora (implica compensar/reducir las dificultades) • factores distractores o que introducen complicación

Investigación y revisión crítica del modelo

Los estudios realizados por los autores y sus colaboradores se han preocupado, entre otros aspectos, de determinar experimentalmente la influencia de los factores que conforman su marco teórico; de este modo, en los primeros estudios (Carlson y Wiedl, 1976, 1978, 1979), se compararon hasta 6 condiciones de aplicación de las matrices progresivas de Raven: (1) instrucción estándar, (2) verbalización durante y después de la solución, (3) verbalización después de la solución, (4) retroalimentación simple, (5) retroalimentación detallada, y (6) retroalimentación detallada más verbalización durante y después de la solución. El test fue administrado a una muestra que osciló entre 108 y 433 alumnos en los diferentes estudios realizados; al comienzo, esta muestra estuvo compuesta por alumnado sin dificultades, pero más tarde se fueron incluyendo niños y niñas con distintos tipos de problemas relacionados con su aprendizaje (déficit

cognitivo, DA, lesión neurológica y deprivación sociocultural). La prueba fue aplicada a niños de 2º y 4º curso en dos tipos de formato: el habitual y mediante puzzles. El análisis de los resultados reveló que, tanto las condiciones de aplicación como el formato de prueba utilizado influían significativamente en los datos obtenidos; en este sentido, las condiciones de aplicación que incluyeron retroalimentación y verbalización durante y después de la prueba se mostraron más eficaces, aunque, como señalan Sternberg y Grigorenko (2003), que realizaron un análisis más detallado de la magnitud de los efectos principales especificados por Carlson y Wiedl en sus análisis, esta magnitud encontrada del efecto de la condición de prueba, a la que antes aludíamos, no era superior al de la repetición simple de la tarea o al de la versión de prueba utilizada.

Posteriormente, los investigadores reestructuraron la situación de prueba basándose en los resultados obtenidos, e incorporaron la verbalización y la retroalimentación a la tarea; de esta forma, durante la actividad y después de la misma, se pedía a los alumnos que describieran la tarea que estaban realizando y que trataran de verbalizar su actividad cognitiva; el evaluador realizaba preguntas para facilitar este tipo de actividad en el sujeto y le proporcionaba información detallada acerca del proceso y de su actuación. Los resultados mostraron de forma significativa una mejora en el rendimiento de los alumnos cuando se procedió a la reestructuración de las condiciones de la prueba; esta mejora afectó al alumnado con déficit cognitivo, con DA, con lesiones neurológicas y con deprivación sociocultural que formaba parte de la muestra (Sternberg y Grigorenko, 2003). También mostraron que las puntuaciones obtenidas en las pruebas reestructuradas predecían mejor el rendimiento escolar que las pruebas aplicadas en su forma tradicional. Como podemos observar, la reestructuración de las condiciones de prueba consistió, en la práctica, en la introducción de una estrategia de intervención de corte metacognitivo, donde en definitiva se le pedía al alumno que pensara en voz alta sobre su propia actividad cognitiva, con el objetivo de favorecer la conciencia sobre el propio proceso de aprendizaje y de que ejerciera un mayor control y una mayor regulación sobre el proceso, debiendo supervisar y evaluar su propia actividad. Sin duda, este aspecto nos parece sumamente interesante, ya que nuestra propuesta considera esencial la introducción de estrategias metacognitivas en el proceso de ED para la obtención de información más valiosa, tanto para determinar el potencial de aprendizaje, como para la orientación a la intervención (Navarro, 2004); por otra parte, no nos es extraño que sea precisamente la introducción de estos procesos en una situación de prueba la que de una manera significativa contribuya a mejorar tanto la información que puede obtenerse de ésta, cuanto el rendimiento del alumnado evaluado con la misma; son varios ya los trabajos que desde nuestro propio contexto han aportado datos en esta misma dirección (Aguilera, 1997; Saldaña, 2001; Montanero, 2000; Moreno y Mora, 2002; Navarro y Mora, 2003b).

Otros análisis efectuados para profundizar en estos hallazgos sobre la mejora de los alumnos cuando se reestructuraban las condiciones de la prueba en el sentido de potenciar la verbalización y la retroalimentación detallada durante y después de la tarea, mostraron que eran precisamente los alumnos que presentaban dificultades en procesos metacognitivos como la planificación, los que mejoraban ostensiblemente su rendimiento gracias a la introducción de las variables mencionadas (Cormier, Carlson y Das, 1990; Kar et. al., 1993); estos resultados apuntaban a que la mayor eficacia del modelo de evaluación propuesto se producía precisamente en la valoración de procesos metacognitivos en sujetos con niveles bajos de rendimiento en estos mismos procesos.

Un aspecto muy importante de los estudios de estos investigadores ha sido su preocupación por las variables de tipo sociopersonal. Ya en uno de sus primeros estudios (Carlson y Wiedl, 1979), citado en Sternberg y Grigorenko (2003), analizaron las correlaciones entre las distintas condiciones de prueba, a las que antes hacíamos mención, y diversos perfiles de personalidad (como la introversión y la ansiedad) y de estilos de aprendizaje (como el estilo reflexivo o impulsivo). Los resultados obtenidos, analizados con una muestra de 203 alumnos de 2º y 230 de 4º, aunque sí mostraron diferentes correlaciones en función de las distintas condiciones de prueba, no indicaron diferencias claras entre la influencia de las variables de personalidad en la aplicación estática de la prueba (condición 1) y la influencia de estas variables en la administración dinámica con retroalimentación detallada y verbalización durante y después de la tarea (condición 6). Estos resultados parecen ser contradictorios con los obtenidos por otros investigadores (Fernández-Ballesteros y Calero, 2000) y también por estos mismos autores en otro estudio (Bethge, Carlson y Wiedl, 1982), en el que estudiando a 72 niños de 3º curso encontraron que la ED reducía significativamente la ansiedad y la actitud negativa ante la situación de prueba frente a una evaluación tradicional. Este dato sí parece encajar mejor con la concepción previa que probablemente compartimos con muchos otros investigadores de la ED con respecto a la relevancia que pueden tener estos modelos en el desarrollo y la mejora de procesos de ajuste personal-social. En este sentido, el análisis cualitativo de la implementación del dispositivo de ED que hemos desarrollado nos ha informado habitualmente de este hecho; sin embargo, posteriormente, cuando abordemos el desarrollo de nuestra investigación, aportaremos otros datos de corte más cuantitativo que matizan los hallazgos previos e introducen la cuestión de la dificultad de acceder y modificar algunos de estos procesos.

Otro de los resultados interesantes del estudio anterior es que se encontró una correlación significativa positiva entre la condición de prueba número 6, es decir, aquella donde la retroalimentación y la verbalización eran más intensas, y la mejora de los niños que presentaban un estilo de aprendizaje impulsivo. Esta misma condición de prueba reducía las diferencias en otro estudio en el que se analizó la ejecución de tres grupos de niños pertenecientes a estratos socioculturales y económicos diferenciados (Dillon y Carlson, 1978: citado en Sternberg y Grigorenko, 2003). En este estudio se usaron 3 condiciones de prueba: aplicación estándar, verbalización y verbalización más retroalimentación; aunque los resultados fueron significativamente distintos entre la aplicación estándar y las otras dos condiciones, las diferencias se reducían también de forma significativa en la condición de verbalización más retroalimentación. Los investigadores concluyen que las diferencias que aparecen entre diversos grupos étnicos y culturales cuando son aplicadas pruebas estáticas para evaluar su funcionamiento cognitivo y sociopersonal, podrían reducirse e incluso eliminarse con la aplicación de pruebas dinámicas; este hallazgo, como recordamos, es congruente con los de Feuerstein y Budoff, para quienes las pruebas dinámicas eran una forma de eliminar el sesgo que suponían diferencias previas en cuanto a las oportunidades habidas de experiencias de aprendizaje mediado, enfrentándose a la evaluación en unas condiciones más equitativas.

Por otra parte, otros estudios sirvieron para reforzar los hallazgos anteriores en los que la modalidad de prueba utilizada incluía procesos de verbalización por parte del alumno (Wiedl y Bethge, 1983); estos procesos de verbalización tenían lugar de forma inducida durante la fase de mediación, en la cual se facilitaba y reforzaba por parte del evaluador el uso de explicaciones en voz alta sobre la propia actividad cognitiva, explicaciones

que podía ser efectuadas a modo de autoinstrucciones o en respuesta a las cuestiones planteadas por el adulto. Carlson, Wiedl y sus colaboradores habían mostrado poco tiempo antes que la verbalización en la aplicación dinámica de los tests de Raven y de Cattell, tenía un efecto de compensación de los problemas que presentaban niños con un estilo de aprendizaje caracterizado por la impulsividad, así como en los niños con elevados niveles de ansiedad y baja motivación (Bethge, Carlson y Wiedl, 1982), y también en niños con déficits en su capacidad de planificación (Cormier, Carlson y Das, 1990). Según los autores del modelo (Carlson y Wiedl, 2000), parece ser que la relevancia que los procesos de verbalización, antes, durante y después de la tarea, podrían tener en la mejora del funcionamiento cognitivo, radica en una posible reestructuración de los procesos de pensamiento en el sentido de optimizar los recursos atencionales, lo que podría afectar positivamente a la consideración de las reglas que deben ser seguidas, las estrategias que deben ser empleadas o la información que debe ser recordada. Otro factor no menos importante que destacan estos investigadores es el que se refiere a que el uso de una modalidad de evaluación e intervención que incluya la verbalización del alumno supone en la práctica la utilización de otro sistema de codificación. En efecto, ante tareas que básicamente son presentadas bajo una modalidad perceptivo-visual (como ocurre en los tests de Raven o Cattell) el uso inducido de procesos de pensamiento en voz alta que planifiquen la propia actuación, describan y analicen la tarea o justifiquen las respuestas, implica una doble y conjunta codificación visual y simbólico-lingüística, lo cual podría explicar una mejora significativa en la actuación.

En un estudio llevado a cabo por Lauth y Wiedl (1985), se propusieron mostrar la influencia de las *características de la tarea* en la ejecución del test de Cattell para un grupo de niños que presentaba DA y otro grupo que no tenía problemas en su aprendizaje. Se presentó uno de los ítems de la prueba consistente en elegir de entre 5 alternativas de respuesta aquella que continuaba una serie de 3 líneas de diferente grosor. Las demandas cognitivas de la tarea fueron modificadas, proporcionándole a los chicos lápices de distinto grosor; primero 3 lápices con distinto grosor que se correspondía con las 3 primeras líneas y después 5 lápices cuyo grosor de línea se correspondía con cada una de las 5 alternativas de respuesta. El resultado fue que los chicos con DA mejoraron significativamente su respuesta, mientras que en los chicos sin DA no se apreciaron cambios con respecto a la ejecución normativa de la tarea. Los autores del estudio concluyeron que la reducción efectuada en la tarea de demandas de tipo lingüístico-conceptual, permitió un procesamiento perceptivo más directo y facilitó el hecho de que pudiera tener lugar un efecto de compensación —en este caso *catalizadora*— de sus déficits cognitivos o metacognitivos, mejorando de esta forma sus resultados.

Carlson y Wiedl han formulado recientemente un modelo de inteligencia con el que se pretende determinar sobre qué elementos debe establecer su actuación la ED y, consecuentemente, sobre cuáles establecer su validez como propuesta de evaluación del funcionamiento cognitivo. Se trata de un modelo comprensivo e interactivo que incluye cinco dimensiones: (a) *capacidades primarias biológicamente determinadas*, (b) *inteligencia fluida* —factor «G»—, (c) *estructuras cognitivas generales*, (d) *estructuras cognitivas específicas*, y (e) una quinta dimensión que estaría constituida por la *cultura*, que desempeñaría un papel mediador progresivo a medida que avanzamos en las dimensiones anteriores (Carlson y Wiedl, 2000).

Las *capacidades primarias biológicamente determinadas* hacen referencia a una serie de estructuras potenciales filogenéticamente desarrolladas que se encontrarían en la base de la adquisición de una variedad de logros considerados universales, como el lenguaje o diversas nociones o logros sensorio-motores descritos por Piaget (Piaget, 1969). Esta dimensión estaría fuera del ámbito de validación y actuación de la ED según estos autores. Por su parte, *la inteligencia fluida* o factor «G», que estos investigadores consideran de forma diferenciada dentro de su modelo, ha ocupado una parte importante de los estudios sobre ED, estudios que han intentado demostrar básicamente que la aplicación dinámica de los tests de factor «G» predecían mejor criterios externos como el rendimiento escolar, que su aplicación estándar. *Las estructuras cognitivas generales* harían referencia a las habilidades cognitivas de orden superior y coincidirían básicamente con la descripción que Vygotski realizó de las funciones mentales superiores, desempeñando un papel de «*mediación psicológica*» en el desarrollo de dominios y funciones cognitivas más específicas. Aquí se encontrarían fundamentalmente las habilidades metacognitivas de autorregulación. Estas estructuras cognitivas generales se encontrarían estrechamente relacionadas con los logros obtenidos en la cuarta dimensión del modelo. *Las estructuras cognitivas específicas* hacen referencia a los dominios específicos como la aritmética, la lectura, la música o la escritura, así como a funciones cognitivas básicas o específicas como la orientación espacial, la coordinación visual o la memoria de trabajo. Es precisamente en estas dos últimas dimensiones del modelo donde los autores consideran más fructífera la aplicación de procesos de ED y la validación de la misma, y de forma especial hacen referencia a las *estructuras cognitivas generales* en su conjunto y a la *memoria de trabajo* (MT) dentro de las estructuras cognitivas específicas. La MT, como hemos tenido oportunidad de constatar durante la revisión de algunos de los modelos aquí presentados, juega un papel fundamental en el funcionamiento cognitivo y en el desarrollo de habilidades y dominios específicos (Swanson, 1995a). Es también un elemento que se ha mostrado eficaz en la discriminación de sujetos que presentan dificultades en su aprendizaje, y su valoración mediante pruebas dinámicas, como el TPC-S (Swanson, 2000), ha contribuido de manera importante a la validación predictiva de la ED con relación a dominios específicos como el lenguaje o las matemáticas.

Con vistas a lograr optimizar los esfuerzos dirigidos a la ED en el sentido antes apuntado, se describen tres pasos necesarios: (1) establecer teórica y empíricamente la relación entre las estructuras cognitivas generales y los dominios y funciones específicas, (2) demostrar la validez predictiva de las pruebas dinámicas sobre los dominios específicos, y (3) desarrollar procesos de intervención educativa basados en la *teoría de las estructuras cognitivas generales* —como instrumentos psicológicos de mediación— y determinar la mayor validez predictiva de las pruebas dinámicas sobre una aproximación evaluativa tradicional.

Numerosos estudios, algunos de los cuales han sido mencionados en este trabajo, se han dirigido a evidenciar el contenido de los dos primeros pasos y, en menor medida, del tercero. Algunos de ellos han tratado de establecer el poder de predicción de la aplicación dinámica sobre la lectura en comparación con pruebas estáticas. Por ejemplo, Hamers, Penning y Guthke (1994) encontraron que los resultados obtenidos con la aplicación dinámica de unas Pruebas de aprendizaje cuyo contenido era relativo al aprendizaje de la lectura y las matemáticas, correlacionaba en mayor medida con los logros obtenidos por el alumnado en esos mismos dominios que con un test de inteligencia o con una Prueba de aprendizaje de tipo general. Lidz, Jepsen y Miller

(1997) también hallaron mejoras significativas en cuanto al nivel de predicción sobre lectura y matemáticas usando y aplicando de forma dinámica unos subtests seleccionados del CAS (Cognitive Assessment System) basado en el modelo PASS (Naglieri y Das, 1987, 1988), en comparación con la aplicación de una prueba estática. Por su parte, Little, Das, Carlson y Yachimowicz (1993) demostraron que habilidades cognitivas generales como la planificación, la atención, así como la capacidad de integración de la información, ejercían un papel mediador significativo en la capacidad de lectura temprana; por otra parte, Carlson y Das (1997) demostraron también que el trabajo desarrollado para mejorar estas habilidades generales, repercutía significativamente en la mejora de la capacidad de lectura, e igualmente en la mejora de las propias habilidades generales.

En definitiva, la intención de los autores es establecer aquellas áreas de funcionamiento cognitivo que sean relevantes desde un punto de vista teórico y empírico en la explicación de dominios específicos; estas serían las áreas en las cuales la validación de la ED tendría un mayor sentido. En este sentido, y tomándolo como referencia, los autores hacen alusión al modelo conceptual propuesto por Beretvas (1999) para los procesos de validación de pruebas dinámicas.

*Modelo conceptual para los procesos de validación de Beretvas (1999),
citado por Carlson y Wiedl (2000)*

- Nivel I Incluye observación e investigación previa, que lleva al desarrollo de constructos teóricos y su operativización (validez de contenido).
- Nivel II Implica juicio de validez estructural. Requiere una exploración de la red conceptual de la que surge el constructo. Los constructos relacionados potencialmente deben ser determinados. Puede ayudar a proporcionar las bases para juicios sobre el Nivel III.
- Nivel III Implica juicio de generabilidad del constructo
- Nivel IV Implica juicio sobre validez de consecuencia. Está relacionado con el contexto y la justificación del uso del test. Debe ser el resultado de juicios pragmáticos combinados con análisis científicos. Si una medida no puede ser justificada por su utilidad práctica es irrelevante.

Envoltiendo estos aspectos se encontrarían, en orden de mayor a menor generalidad, *la justificación para el uso del test* y *el contenido del test*, es decir qué evalúa. Requeriría una justificación clara y científica de los constructos desarrollados, procedimientos de evaluación empleados y dispositivos de evaluación usados. Requeriría igualmente evidencias de su utilidad práctica, lo que a su vez exige investigación básica y aplicada.

En síntesis, el modelo de Prueba de los límites ha mostrado o apuntado hacia la idoneidad de intervenir sobre procesos de autorregulación y conocimiento metacognitivo para mejorar el rendimiento de aquellos alumnos que muestran precisamente un bajo nivel de funcionamiento metacognitivo. También ha mostrado la relevancia de aspectos relativos a la personalidad en el rendimiento diferencial ante pruebas estáticas y dinámicas. Su propuesta, como hemos visto, consiste en la reestructuración de las pruebas convencionales para introducir aspectos que se han revelado cruciales en la mejora de la ejecución, como son la verbalización durante y después de la tarea y la retroalimentación detallada del evaluador. Estas condiciones de

prueba no requieren por tanto un modelo clásico pretest/ mediación /postest, lo que minimiza también algunos problemas metodológicos como la medición y fiabilidad del cambio entre pretest y postest. Probablemente, la mayor objeción que podría hacerse a este modelo radica en la dificultad para el establecimiento de comparaciones individuales, ya que la metodología de estudio arroja resultados grupales que comparan diversos modos de intervenir (procedimiento estándar, verbalización, retroalimentación, etc.) con el objetivo de determinar la modalidad de intervención óptima para cada grupo (Sternberg y Grigorenko, 2003).

Otro aspecto que puede ser interpretado como una limitación de este modelo lo constituye el hecho de que los resultados obtenidos de sus estudios presentan a menudo la interacción de diversos factores en la explicación; de esta forma, la modalidad de retroalimentación detallada junto con la verbalización parece ser óptima para alumnos que cumplan varias condiciones: bajo funcionamiento cognitivo medido con pruebas de aptitud, niveles elevados de ansiedad, encontrarse en una determinada etapa del desarrollo, y ante tareas que requieran un alto nivel de procesamiento cognitivo. Lógicamente, este aspecto tiene una doble lectura; por una parte, limita de forma significativa la población beneficiaria de la propuesta, pero, por otra, define de forma explícita y concreta esta población para la cual los beneficios diferenciales de afrontar una situación de aprendizaje o evaluación que se caracterice por desarrollarse en la modalidad mencionada, son muy importantes con respecto a un proceso estándar.

Por último, otro asunto muy importante es el que hace referencia a la validez predictiva del modelo. En este caso, los estudios revelan que cuando la modalidad de prueba consiste en la conjunción de la retroalimentación detallada y la verbalización, las puntuaciones obtenidas predicen mejor el rendimiento en Lengua y Matemáticas. En este sentido, también se ha mostrado la relevancia de que la modalidad de prueba, en definitiva la metodología de evaluación adoptada, y la metodología de enseñanza, guarden una estrecha relación; en efecto, en un estudio de Wiedl y Herring (1978), el rendimiento escolar en matemáticas fue predicho mejor por las puntuaciones obtenidas de la aplicación dinámica de los tests de Cattell y Raven, que utilizaba la modalidad de verbalización y retroalimentación detallada, cuando los procedimientos de enseñanza de ese área se correspondieron con la metodología citada —esta metodología se correspondía en gran medida con los principios de la *enseñanza adaptativa*, que hemos visto con anterioridad cuando analizábamos el modelo propuesto por Guthke—. Esto remite a la *validez ecológica* de las pruebas de evaluación y de la metodología empleada por éstas, como uno de los aspectos cruciales de cualquier propuesta que quiera ser útil en un contexto determinado.

La Batería de Evaluación Dinámica de Procesos Metacognitivos de D. Saldaña

Bases teóricas y metodológicas

La consideración y constatación de la relevancia de los procesos metacognitivos en la explicación de la conducta inteligente o adaptativa de los sujetos, así como para la intervención educativa, y muy especialmente, aquella dirigida a personas con necesidades educativas especiales, constituyen la base fundamental sobre la que se asienta el modelo construido por Saldaña (Saldaña, 2001). Se pretende con esta propuesta valorar los procesos de autorregulación, así como la presencia de

metaconocimientos en una población compuesta por sujetos que presentan déficit cognitivo severo, mediante la creación de un espacio compartido entre el evaluador y el sujeto, donde sea posible provocar experiencias de discusión consciente de la propia actividad con el fin de obtener información adicional sobre el funcionamiento cognitivo. Por otra parte, el análisis que realiza el autor de los distintos métodos de evaluación de la metacognición, revela la necesidad de contar con instrumentos capaces de captar y analizar ciertas habilidades que pueden hallarse ocultas debido a la presencia de dificultades o limitaciones en otras; por esto se considera necesaria la aplicación de un modelo de ED donde se facilite la aparición de aquellas habilidades o procesos incipientes del sujeto, que no son observables en un primer momento mediante la aplicación de una prueba convencional. Es precisamente la «simbiosis» entre evaluación e intervención, propia de la ED, la que permitirá obtener información valiosa sobre las habilidades plenamente desarrolladas y también sobre aquellas que se encuentran en proceso de maduración o que no afloran por la existencia de otras limitaciones.

En función de las medidas criterio utilizadas, este trabajo se encontraría más cerca del modelo de Brown y Campione, y su método de las sugerencias graduadas, que del modelo de Feuerstein. Estas medidas se basan en la ayuda necesitada por los sujetos, ya que se considera esta opción más útil y cercana a la experiencia educativa de enseñanza y aprendizaje. Al mismo tiempo, se pretende encontrar un cierto equilibrio entre la naturaleza cualitativa y clínica de la propuesta, y la necesidad de utilizar ciertos parámetros que contribuyan a homogeneizar, al menos en parte, el proceso de evaluación con el fin de obtener cierta fiabilidad. Entre los objetivos del estudio realizado por Saldaña se encuentra el de generar una estrategia metodológica orientada a la evaluación de los procesos metacognitivos en la población mencionada anteriormente, al tiempo que sentar las bases para el desarrollo de nuevos instrumentos de aplicación dinámica. Se esperaba que el procedimiento utilizado contribuyera a incrementar la validez predictiva de los procesos de evaluación interactiva; para esto, un aspecto relevante de las propuestas de este investigador lo constituye la intención de establecer relaciones entre la metacognición valorada en interacción y la que manifiestan los sujetos en su conducta adaptativa. Esta capacidad de predicción del modelo utilizado se esperaba encontrar también con relación a aprendizajes desarrollados en una situación de intervención prolongada; tal valoración se realizaría con el criterio otorgado por la participación de los sujetos en el programa de enriquecimiento cognitivo *Comprender y Transformar* (CyT) (Mora, 1991, 1998), de orientación metacognitiva.

La Batería de Evaluación Dinámica propuesta consta de 5 actividades que comparten los mismos parámetros para la valoración, así como también las pautas específicas para el proceso evaluador. Estas actividades son: *Atención, Memoria, Matrices Progresivas, Mastermind e Interacción Social*, y con ellas se pretende acceder a la valoración de los metaconocimientos y los procesos de autorregulación. En cuanto a los metaconocimientos, se distinguen en este estudio 3 categorías: (a) *metaconocimientos sobre posibilidades y limitaciones*, que haría referencia a la comprensión que el sujeto tiene de sus propias habilidades con relación a la tarea; (b) *metaconocimientos sobre las dificultades y características propias de la tarea*, tales como su estructura, lo que diferenciaría, por ejemplo, modalidades más complejas de otras más sencillas; y (c) *metaconocimiento estratégico*, que haría referencia a los planes de acción, algoritmos, posibles vías de solución o heurísticos de los que dispone el sujeto para la resolución del problema. Por lo que respecta a los procesos de autorregulación se distinguen: (a) los

procesos de planificación, dentro de los cuales se insertan la identificación del problema y de los objetivos de la tarea, y el desarrollo y/o selección de estrategias y operaciones para la solución del problema; (b) los *procesos de supervisión*, que integran el mantenimiento de objetivos y la valoración de su consecución durante la tarea, así como el control sobre la adecuación de la aplicación de las estrategias; y (c) los *procesos de evaluación*, que hacen referencia a la valoración del logro y grado de éxito alcanzado con relación a las metas definidas, y a la evaluación de la adecuación de la estrategia adoptada.

En el proceso de mediación se utilizan 5 modalidades de interacción, cuatro de las cuales son aplicadas gradualmente en función de la respuesta del sujeto, siendo aplicada de forma general la modalidad restante. En la lógica de las sugerencias graduadas, cuando el sujeto realiza con éxito una tarea, no es necesario seguir avanzando en la aplicación de modalidades que impliquen una mayor asistencia del proceso. Estas modalidades son:

- *Observación con o sin verbalización*. Se pide al sujeto que realice la tarea y, en función de su capacidad comunicativa, que verbalice los pensamientos que acompañan su acción.
- *Asistencia mediante apoyo puntual*, en función de las dificultades observadas en la realización de la tarea.
- *Apoyo global en la tarea*. Consiste en un apoyo generalizado en procesos de autorregulación, mediante la administración de instrucciones específicas que conducen al sujeto hacia la solución correcta del problema.
- *Modelado*. Es utilizado ante casos en los que no se produce una respuesta positiva del sujeto a pesar del apoyo global del evaluador; en estos casos, el procedimiento que se adopta es la ejecución de la actividad por parte del evaluador con la intención de que sea seguida y replicada posteriormente por el sujeto.
- *Entrevista*. Esta última modalidad de interacción constituye en realidad un procedimiento complementario de los anteriores y consiste en preguntar al sujeto, antes o después de la tarea, acerca de sus conocimientos sobre la misma y sobre su resolución, características y dificultades de la actividad, estrategias adoptadas, etc. Esta modalidad no es aplicada de forma gradual, sino que es utilizada en todas las tareas en distintos momentos de la misma.

La medida del constructo evaluado en cada una de las cinco actividades propuestas viene dada por la interacción entre la modalidad de interacción utilizada y la habilidad metacognitiva valorada. De esta forma, para cada dimensión metacognitiva utilizada se establece la modalidad de apoyo requerida, configurando un cierto perfil de la competencia metacognitiva del sujeto en función de una graduación que va desde la ausencia a la presencia de autorregulación: (1) el apoyo no tiene impacto sobre la ejecución, (2) apoyo al límite de las posibilidades del sujeto, (3) necesidad de heterorregulación permanente, (4) necesidad de heterorregulación frecuente, (5) autorregulación. Es esta configuración la que permite la valoración cualitativa de la prueba, mediante la asignación de 1 a 5 puntos en función de la capacidad manifiesta de autorregulación del sujeto en cada una de las tareas.

Investigación y revisión crítica del modelo

El estudio presentado (Saldaña, 2001) fue realizado con una muestra de 29 sujetos, de los cuales 16 constituyeron el GE y 13 el GC. Los integrantes del GE fueron seleccionados entre estudiantes de un Taller ocupacional especial y un colegio concertado de Educación Especial en Sevilla (Andalucía). Las características de esta muestra eran: poseer un nivel mínimo de comunicación que permitiera la interacción con el evaluador, ausencia de problemas graves de conducta, y un nivel de funcionamiento cognitivo significativamente bajo (CI medio en Raven = 25,53; DT = 13,33). La edad media de la muestra era de 19,6 años, siendo el grupo equivalente en cuanto a la presencia de ambos sexos. Por su parte, el GC fue seleccionado de entre dos Asociaciones juveniles de Jerez de la Frontera, en Cádiz (Andalucía), y las características básicas eran la ausencia de problemas significativos en su desarrollo cognitivo y estar escolarizados en un nivel superior al de ESO o equivalente. La edad media en este grupo, que estaba compuesto por 8 chicos y 5 chicas, era de 18,5 años y el CI medio en la prueba de Raven fue de 98,37, siendo su DT de 9,25.

Uno de los aspectos más novedosos y relevantes de la propuesta de Saldaña, lo constituye la utilización de actividades lúdicas con base en materiales y juegos conocidos y significativos para la población a la que se dirigía el estudio y que han sido seleccionados y preparados específicamente para servir a los objetivos de la evaluación. Entre estas actividades se encuentran, por ejemplo, el juego Mastermind o la búsqueda de Wally. En el siguiente cuadro, que tomamos del trabajo de Saldaña, recogemos estas actividades junto con el objetivo que cumplen dentro de la propuesta.

*Cuadro 21. Actividades y procesos que evalúan en la Batería de Evaluación Dinámica.
Tomado de Saldaña (2001)*

Evaluación centrada en aspectos de metacognición y control metacognitivo	Actividad
Atención	<i>¿Dónde está Wally?</i>
Memoria	<i>Lista de la compra de 26 elementos</i>
Razonamiento inductivo/medida de la inteligencia	<i>Matrices Progresivas estándar de Raven</i>
Resolución de problemas/Habilidades de pensamiento hipotético-deductivo	<i>Mastermind (versión adaptada)</i>
Resolución de problema interpersonal	<i>Actividad simulada de una situación problemática de interacción social mediante la utilización de muñecos</i>

Nos parece, así mismo, de especial interés, la utilización en la Batería de evaluación propuesta de actividades que traten de conectar el proceso de evaluación con el desempeño metacognitivo en situaciones reales o simuladas en las que sea preciso poner en práctica estas habilidades de una manera adaptativa. Estas actividades han sido utilizadas ya en nuestro contexto (Aguilera, 1997; Mora, 1998), así como en otros (Rogoff, 1993), en el marco de procesos de evaluación del pensamiento en interacción y

de programas de enriquecimiento cognitivo, y constituyen a nuestro parecer una aproximación relevante al estudio contextualizado de las habilidades de pensamiento. En este sentido, se utiliza en el estudio que ahora revisamos una escala de conducta adaptativa, el ICAP (*Inventario para la Planificación de Servicios y Programación Individual*) (Bruininks et. al., 1986/1993), diseñada especialmente para la evaluación de personas con discapacidad. La utilización de esta escala ofrece al autor una información complementaria a la obtenida a través de la Batería de Evaluación Dinámica con relación a factores que han ido siendo considerados progresivamente claves en la definición y la conceptualización de la deficiencia mental; estos factores hacen referencia básicamente al funcionamiento sociopersonal autónomo, a las relaciones interpersonales y a la responsabilidad social (Montero, 1993), citado en Saldaña (2001). También como medida complementaria, se utilizó un tercer instrumento de evaluación incluido en una versión del programa CyT, diseñada para sujetos con daño cerebral (Moreno y Mora, 2002; Mora y Moreno, 1998; Moreno, Pastor y Piñero, 1995). Esta versión del programa mantiene la naturaleza metacognitiva y de interacción para la estimulación de las habilidades de pensamiento, pero a diferencia de la versión original en la que las actividades tienen lugar en grupo, se desarrolla en un entorno informatizado y mediante la interacción entre evaluador y sujeto. El instrumento de valoración es la *Matriz de Conductas Cognitivas Enriquecidas* (Mora y Mora Merchán, 1995). Esta matriz se compone de cuatro subescalas: Gestión de la Información, Aplicación a la Solución de Problemas, Autocontrol/Control Metacognitivo y Ajuste Personal-Social.

Por lo que respecta a los resultados generales del estudio, éstos indican de entrada diferencias muy significativas entre los grupos experimental y control en cuanto a los indicadores metacognitivos en todas las tareas utilizadas. Dado que las puntuaciones otorgadas en la Batería de Evaluación Dinámica, así como en la matriz de conductas cognitivas enriquecidas lo eran con base en las observaciones de la puesta en práctica de indicadores de conducta de los sujetos, se procedió a evaluar la fiabilidad de estas valoraciones mediante procesos de acuerdo inter-jueces; estos resultados, en general, fueron satisfactorios, tanto por lo que respecta al acuerdo en la definición de los indicadores de conductas metacognitivas, como en la fiabilidad de las observaciones según la tarea de evaluación.

Una de las críticas habituales a los procedimientos convencionales de evaluación del funcionamiento cognitivo es que no ofrecen una discriminación fina del segmento de población con bajas puntuaciones; la ED, como hemos tenido oportunidad de ver a lo largo de la revisión de las diferentes propuestas aquí recogidas, ha tratado de mostrar en numerosas ocasiones su validez para este fin. Entre sus objetivos, el estudio de Saldaña también pretendía detectar distintos grados de desarrollo metacognitivo en los sujetos con deficiencia mental evaluados. La hipótesis era, también en este caso, que el empleo de la cooperación y la asistencia al sujeto evaluado facilitaría una información adicional a la ofrecida en un entorno de ejecución independiente y permitiría establecer diferencias de grado entre los sujetos. Se pretendía así mismo que esta distinción cualitativa fuera útil con relación a otros criterios de utilidad clínica o teórica. Entre estos criterios, fueron estudiados principalmente: la conducta adaptativa y las habilidades metacognitivas desplegadas en el programa de enriquecimiento cognitivo.

Tomando como referencia la conducta adaptativa, evaluada con el ICAP, se procedió a discriminar entre sujetos con bajas, medias y altas puntuaciones y, así mismo, esta

distinción se realizó con los resultados de la Batería de ED; los resultados mostraron que los sujetos con bajas puntuaciones metacognitivas tendían a obtener bajas o medias puntuaciones en el ICAP; los sujetos con una puntuación media metacognitiva, se situaban también en el nivel medio de puntuaciones ICAP; y los que puntuaron alto en metacognición, también presentaron un nivel de conducta adaptativa media o alta. La correlación fue significativa (0,5). Esta capacidad de discriminación de la ED se mostró superior a la obtenida con las medidas de CI, con las que no se alcanzó el nivel de diferenciación en grupos de alta, media y baja calidad de adaptación. El investigador encontró que no todas las tareas de la batería tenían el mismo peso a la hora de su contribución a la correlación con la conducta adaptativa; de forma aparentemente sorprendente, las puntuaciones obtenidas en las tareas de *Memoria de la lista de la compra* y de *Interacción social*, es decir, aquellas que guardaban una más estrecha relación con actividades cotidianas significativas para los sujetos, introducían confusión en los resultados y hacían decrecer la correlación, mientras que las tareas de *Atención*, *Mastermind* y, sobre todo, *Matrices progresivas*, permitían establecer de un modo más claro y significativo el nivel de adaptación de los sujetos. El autor alude a la posible dificultad relativa de algunas actividades o a la mayor facilidad de estructuración de la interacción en otras, como variables que pudieron tener más peso en los resultados que la aparente validez ecológica de las actividades.

Tomando esta vez como referencia las habilidades metacognitivas de los sujetos, se pretendía valorar la validez predictiva de la Batería de ED en la actuación de los chicos en el programa CyT, valorada ésta en términos metacognitivos. La implementación del programa tuvo lugar durante 17 sesiones como promedio, y previamente los sujetos se habían enfrentado sin el apoyo del evaluador a las tareas que son trabajadas en las unidades del programa, lo que permitió obtener unas puntuaciones a modo de pretest sobre su actuación. Se procedió, al igual que con el Inventario de conductas adaptativas, a establecer tres grupos en función de las puntuaciones obtenidas en las sesiones del programa utilizando la matriz de conductas cognitivas enriquecidas. Los resultados mostraron que los sujetos con puntuaciones más bajas en la subescala Control Metacognitivo de la matriz, obtenían también puntuaciones bajas en la batería dinámica; igualmente, los sujetos con puntuaciones altas en Control Metacognitivo, puntuaban alto en la batería. Aunque la distribución de las puntuaciones medias no estaba tan clara, la correlación fue muy significativa (0,64). En cuanto a las otras subescalas de la matriz, se obtuvo que las puntuaciones ordenadas en los tres grupos en Gestión de la Información correlacionaban significativamente con las puntuaciones dinámicas; sin embargo, ni las puntuaciones en Aplicación a la Solución de Problemas, ni las de Ajuste Personal-Social parecían tener relación con las puntuaciones de la batería a tenor de los resultados. Cuando se incorporó a la puntuación que permitía la discriminación de los tres grupos, mediante un análisis de regresión, el grado de mejora experimentado durante las sesiones de intervención, tampoco se obtuvo relación significativa entre el Ajuste Personal-Social y las puntuaciones dinámicas de la batería. Sí se fortalecieron en cambio las relaciones entre la subescala de Control Metacognitivo y la valoración dinámica. A pesar de las limitaciones encontradas en el valor predictivo de la Batería de ED, sobre todo con respecto a los procesos emocionales y motivacionales, sí parece claro que ofrece un valor predictivo adicional y relevante del funcionamiento metacognitivo en una situación de aprendizaje prolongado.

Una de las implicaciones más relevantes que el estudio destaca con respecto a la metacognición, tiene relación con la posible consideración de ésta como un elemento

general, que trasciende distintas situaciones y dominios, o, por el contrario, como un elemento sujeto y dependiente del dominio específico en que tenga lugar. Los resultados muestran lo que el autor denomina una *opción mixta*; a saber, parece haber una influencia decisiva de la tarea específica en los resultados metacognitivos, pero, a su vez, también puede deducirse de los datos que las puntuaciones obtenidas en las distintas pruebas no son del todo independientes, por lo que puede interpretarse una cierta transituacionalidad de la capacidad metacognitiva. Por otra parte, los datos analizados apuntan también a que se produciría una mayor relación entre la valoración dinámica y las puntuaciones de las sesiones del programa que incluían la mejora progresiva de los sujetos; estos datos parecen reforzar la idea de que la ED aportaría información más valiosa sobre la *construcción* de unas determinadas funciones —en este caso funciones metacognitivas— que sobre una medida de estas funciones ya «cristalizada».

Una conclusión que nos parece muy relevante con respecto a la ED de la metacognición, tiene que ver con el hecho de que la evaluación de la metacognición en una situación compartida, mediante la creación de una ZDP, incrementaría la posibilidad de valoración de capacidades ya establecidas junto a aquellas que se encuentran en proceso de maduración; estos dos tipos de funciones cognitivas, que encontraríamos en una situación de ED, no serían independientes, sino que se encontrarían estrechamente relacionadas, y la respuesta adaptativa del sujeto en una situación problemática y en un momento determinado, reflejaría la conjunción de ambos estados de sus funciones y no solo la actuación de habilidades ya cristalizadas.

Abordamos a continuación otros aspectos críticos relacionados con varios elementos relevantes del estudio revisado; uno de ellos, comentado también por el autor, es que debido al exiguo tamaño de la muestra, en conjunción con el hecho de que el nivel de profundización en la definición y valoración cualitativa, tanto de las funciones cognitivas como del desarrollo de los sujetos, no ha sido excesivo, es aconsejable tomar con cautela los datos obtenidos, sobre todo en lo referente a su posible generalización. Por otra parte, en cuanto a las tareas utilizadas y los procesos que básicamente eran valorados en ellas: atención, memoria, razonamiento inductivo, habilidades de pensamiento hipotético-deductivo y capacidad de resolución de problemas interpersonales; nos parece relevante el hecho de que su elección, al menos en parte, haya sido originada por un análisis del contexto en relación a la población destinataria del estudio, en este caso personas con déficit cognitivo, contemplando una serie de funciones cognitivas o de adaptación social consideradas de especial importancia en su desarrollo cognitivo y sociopersonal. Esto, sin duda, se relaciona con la posibilidad de incrementar la validez ecológica del estudio. El hecho de haber utilizado además criterios externos relativos a la conducta de adaptación social y personal refuerza este aspecto positivo del estudio. Por tanto, la contextualización de los instrumentos de evaluación, aspecto clave para nosotros en un proceso de ED que quiera contribuir a explicar y optimizar un determinado proceso, tendría lugar en este estudio en la medida en que las tareas específicas y los procesos que se pretenden evaluar con ellas guarden relación con la actividad que habitualmente deben desempeñar los sujetos, mostrándose valiosas para el desempeño de la misma. Este aspecto, en cambio, como recoge también el autor y hemos analizado en alguno de los modelos presentados anteriormente, puede conllevar ciertos problemas metodológicos, ya que, si no se definen y delimitan, o controlan adecuadamente los niveles de partida de los sujetos cuando son utilizadas tareas con un alto componente familiar para los mismos, es posible, como así ocurre en

la investigación de Saldaña, que no estemos seguros de si distintos niveles previos de metaconocimientos sobre contenidos o procedimientos con relación a las distintas tareas, han contribuido a la determinación de las diferencias obtenidas por los sujetos.

Por último, otro elemento a considerar, del que también hablaremos cuando analicemos nuestro dispositivo de ED, es el que hace referencia a la medida en que un instrumento de estas características es generalizable, atendiendo esta vez a las necesidades de formación y tiempo que conlleva. Estamos de acuerdo con Saldaña en que, tanto los conocimientos necesarios sobre procesos de evaluación en interacción, como los conocimientos sobre procesos metacognitivos, así como la combinación de estos, es decir, la ED de la metacognición, parecen ser elementos clave para abordar un proceso de evaluación como el que se recoge en el estudio presentado; y que estos elementos forman —en contadas ocasiones— o deberían formar parte del proceso de formación inicial y permanente de los psicólogos educativos, aunque probablemente también compartimos con él la constatación de que es mucho el camino que queda por recorrer en este sentido. Por otra parte, es preciso contar con el tiempo necesario para la implementación de una batería como esta, lo cual no parece fácil en la práctica profesional. De todo ello podrían desprenderse dos conclusiones: de una parte, la necesidad, compartida con la mayoría de las propuestas dinámicas, de mejorar el proceso de evaluación mediante la búsqueda de equilibrio entre la valoración de procesos en contextos de interacción y cooperación, y el análisis de la actividad, que nos lleve hacia una optimización del esfuerzo y el tiempo empleados; de otra parte, es también necesario introducir con redoblado esfuerzo la reflexión en la comunidad educativa y profesional sobre los objetivos de la evaluación y sobre si éstos se cumplen verdaderamente con determinadas prácticas que conllevan escaso tiempo y nivel de formación.

En síntesis, este modelo pretende ofrecer la posibilidad de evaluar dinámicamente los procesos metacognitivos y esto constituye ya un aspecto de enorme importancia para nosotros y, en general, para el campo de estudio de la ED. El hecho de que consideremos estos procesos como clave en la explicación del desarrollo de la competencia en numerosos dominios específicos y funciones cognitivas específicas, así como en la presencia y evolución de las DA o la deficiencia mental, hace que prestemos una especial atención al desarrollo de esta propuesta y a la prueba utilizada, así como a futuros análisis que la desarrollen o amplíen sus resultados.

Pruebas dinámicas para revelar el potencial de aprendizaje oculto. Los estudios de R. Sternberg y sus colaboradores

Bases teóricas y metodológicas

Aunque Sternberg no ha propuesto de forma concreta un procedimiento de ED, sí ha abordado, tal como vimos al comienzo de este mismo capítulo, el estudio de las aptitudes intelectuales y socioemocionales como *formas de experiencia en desarrollo* en contraposición a la consideración de éstas como entidades que nos vienen dadas y que se caracterizarían por su carácter esencialmente estable y estático. Es este convencimiento que muestran Sternberg y sus colaboradores (Sternberg y Grigorenko, 2003) acerca de la naturaleza dinámica y experiencial de las aptitudes que conforman el bagaje de las funciones cognitivas, lo que conduce de una forma lógica a la consideración de los procedimientos dinámicos de evaluación como la metodología más

coherente y válida para aproximarse a la valoración del estado dinámico de las funciones en desarrollo de un sujeto.

Investigación y revisión crítica del modelo

Recogemos de forma más breve la referencia a unos estudios realizados por Sternberg y sus colaboradores, en los que se trata de poner en práctica la teoría de las *aptitudes como formas de experiencia en desarrollo* en el ámbito práctico de la ED. En un estudio reciente, citado en Sternberg y Grigorenko (2003), que, por otra parte, recoge ideas similares a las que fundamentaban algunas de las investigaciones expuestas anteriormente, Sternberg y diversos colaboradores (Sternberg, Grigorenko, Ngorosho, Tantufuye, Mbise, Nokes, Jukes y Bundy, 2001), se propusieron demostrar la mayor validez de la ED para determinar el potencial de aprendizaje y el desarrollo de las aptitudes de niños de países empobrecidos, que de otra forma podrían quedar ocultos. Se idearon tres tipos de tareas que fueron administradas al alumnado de escuelas rurales de Tanzania. Las tareas eran similares a las que se encuentran en pruebas convencionales de aptitud intelectual; se pretendía de esta forma obtener puntuaciones estáticas y dinámicas que permitieran comparar ambos enfoques de evaluación.

Se aplicaron dinámicamente tres actividades, siguiendo un modelo pretest/ mediación /postest: *silogismos, clasificación y veinte preguntas*, a un GE compuesto por 358 alumnos y alumnas de 2º a 5º curso, con edades comprendidas entre los 11 y los 13 años. Otros 100 alumnos de igual edad y condición social formaron el GC. También se aplicaron otras pruebas de aptitud y de rendimiento escolar en lectura y matemáticas, que sirvieron de referencia en el estudio. Por otra parte, el pretest y el postest estaban compuestos por tareas similares, aunque no idénticas, y la intervención se basaba en la enseñanza de estrategias que permitían mejorar la ejecución en las tareas. En general, los resultados de distintos análisis mostraron diferencias significativas entre el pretest y el postest en las tres tareas aplicadas.

Analizando las correlaciones entre las puntuaciones pretest y postest para las tres tareas, éstas fueron significativamente superiores para el GC, lo que indicaba un efecto de la intervención en el GE sobre el rendimiento en el postest. Los investigadores concluyen a partir de estos datos que cualquier conclusión que se extrajera contando exclusivamente con la información proporcionada por las puntuaciones pretest, debería considerarse cuanto menos dudosa, ya que las puntuaciones experimentaron un cambio significativo incluso con una cantidad muy pequeña de intervención. También fueron significativamente mayores las correlaciones del postest con las medidas de las pruebas de aptitud que fueron administradas como referencia. Estos resultados mostraron que las puntuaciones postest predecían mejor las medidas de aptitud que las pretest. Con respecto al rendimiento escolar en lectura y matemáticas, las puntuaciones en los postest presentaron correlaciones también significativamente mayores con estas medidas que las puntuaciones de los pretest.

Desde la perspectiva teórica formulada por Sternberg (1985, 1997a, 1998, 1999a), relativa a las *aptitudes como formas de experiencia en desarrollo*, y desde otras concepciones que pueden ser consideradas precursoras de la misma (Vygotski, 1934/1995, 1979/1996; Feuerstein, 1979, 1991; Budoff, 1975), se enfatiza que muchos niños de entornos social y culturalmente desfavorecidos, reflejan niveles altos de capacidad cuando éstas son medidas como aptitudes que se hallan en continuo

desarrollo y no exclusivamente ya desarrolladas. Los investigadores del estudio que hemos expuesto observaron que los alumnos y alumnas participantes mostraban aptitudes intelectuales que no eran recogidas por la aplicación estática de las mismas pruebas. Estos autores piensan que las pruebas dinámicas que adoptan un modelo pretest/ mediación /postest podrían convertirse, de esta forma, en un procedimiento complementario y no sustitutivo de los procedimientos estáticos, al incorporar en su diseño una prueba pretest aplicada convencionalmente.

En síntesis, qué duda cabe acerca de que el modelo teórico propuesto por Sternberg con relación a la consideración de todas las aptitudes humanas como formas de experiencia en desarrollo, fundamentado también en su teoría triárquica sobre la inteligencia, podría contribuir de manera significativa al desarrollo de procedimientos dinámicos de evaluación más comprensivos y complejos. Así mismo, la preocupación manifestada por él y sus colaboradores sobre la fiabilidad y la validez de las pruebas dinámicas desarrolladas, hacen pensar que los estudios que se deriven del entorno científico de este autor, aporten ciertas garantías sobre estas cuestiones metodológicas. Por otra parte, aún se han realizado pocas investigaciones que hayan partido claramente de los postulados teóricos de Sternberg, por lo que habrá que seguir atentamente la evolución de estos estudios. La consideración de todas las aptitudes como formas de experiencia en desarrollo, así como la contemplación de la ED como el procedimiento que permitiría aproximarse a una valoración más real y comprensiva de las funciones de un sujeto, debería conducir en buena lógica al desarrollo de modelos dinámicos de evaluación, no solo de la inteligencia como constructo o de funciones cognitivas generales, sino de procesos implicados en el desarrollo de dominios específicos más relacionados con la actividad educativa escolar. Hasta el momento, como decíamos más arriba, no son muchos los estudios sobre ED realizados desde este modelo, pero parece muy prometedor que algunos de ellos hayan tenido como objeto la comprensión de textos en una segunda lengua (Grigorenko, Sternberg y Ehrman, 2000; Sternberg y Grigorenko, 2003).

La Evaluación Dinámica de la competencia lingüística en una segunda lengua

La evaluación dinámica está tratando de demostrar su validez para otorgar un «plus» de información en ámbitos distintos a los que tradicionalmente ha dedicado su atención, ampliando su campo de acción e investigación para adentrarse en la esfera de dominios específicos como por ejemplo el aprendizaje de una segunda lengua. Los estudios de Grigorenko, Sternberg y Ehrman (2000), llevados a cabo en Estados Unidos, sobre el supuesto de que la adquisición de un idioma extranjero es también una forma de *experiencia en desarrollo*, y los de Garb y Kozulin (1998, 1999) en Israel, abogan por la aplicación de pruebas dinámicas para evaluar la competencia lingüística en el aprendizaje de una segunda lengua. Como expresan Kozulin y Garb (2001), la ED debería proporcionar información sobre la capacidad de aprender, activar y usar de forma efectiva, estrategias de comprensión de textos en una segunda lengua.

La Evaluación Dinámica de la Comprensión de Textos en una Lengua Extranjera, de A. Kozulin y E. Garb

En un interesante estudio llevado a cabo por Garb y Kozulin (1999), se realizó una aplicación dinámica de una prueba de evaluación de la competencia lingüística en una

lengua extranjera. El test fue adaptado para que la valoración fuera exclusivamente sobre comprensión lectora, eliminando aquellos ítems que valoraban sólo vocabulario, lo cual remitía de forma clara a los conocimientos y experiencias previas del alumnado. Los ítems válidos de la prueba fueron analizados en un doble sentido: (a) qué conocimientos era preciso tener para poder resolver las tareas, y (b) qué estrategias eran necesarias para su resolución. La muestra estaba compuesta por 23 estudiantes con bajo nivel académico, de edades comprendidas entre 18 y 25 años de un Centro de adultos de un área rural de Israel. Por otra parte, el formato usado para la investigación fue el de pretest/ mediación /postest, consistiendo la mediación básicamente en la revisión conjunta de la prueba pretest con el profesor, analizando los errores cometidos y abordando las estrategias necesarias para su corrección, así como dónde y cómo podían ser transferidas estas estrategias a otras actividades. Más concretamente, en la primera parte de esta fase de mediación, durante dos sesiones, se instruyó en la manipulación de estructuras sintácticas y gramaticales (estructuras de sujeto-verbo-complemento, etc.), proporcionando además información básica de tipo léxico y gramatical, sobre verbos auxiliares, oraciones interrogativas, negativas, etc. La segunda parte de la mediación incluía preguntas de comprensión sobre la base de cuatro textos. Se instruía también en estrategias de tipo estructural y en la activación de conocimientos previos que elicitaran la construcción del significado del texto y de las preguntas formuladas. Estas sesiones de mediación duraron 50 minutos.

Los resultados mostraron una mejora del grupo por encima de una desviación típica; la puntuación de mejora (en este caso, puntuación pretest *menos* puntuación postest) estaba negativamente relacionada con el pretest, lo que indicaba, según los autores, que la puntuación obtenida en la fase inicial mostraba el nivel de logro actual, pero no reflejaba el potencial de aprendizaje de los estudiantes. Se constató que alumnos y alumnas con idénticas puntuaciones pretest, puntuaron de forma muy desigual en el postest; igualmente, alumnos con bajas puntuaciones en la fase pretest, puntuaron mejor en la fase de mediación que otros que habían obtenido mejores puntuaciones iniciales, indicando mejor su potencial de aprendizaje. La ED proporcionaría de esta forma una información más detallada y valiosa sobre la capacidad de aprendizaje y sobre las diferentes necesidades educativas.

Los autores concluyen que la ED ofrece información útil, no solo en el campo de las funciones cognitivas generales, sino también con relación a dominios específicos relacionados con áreas académicas como por ejemplo la lectura. Los autores también hacen una reflexión sobre las posibles limitaciones del estudio en cuanto a su generalización; es preciso en este sentido hacer notar el número bajo de estudiantes que participaron en el estudio, y, aún más importante, la ausencia de grupo control que, sin duda, hubiera enriquecido los resultados. Por otra parte, Kozulin y Garb aluden a que los resultados en un proceso de ED reflejan, no solo el potencial de aprendizaje del alumnado participante, sino también la *calidad y cualidad de la mediación* realizada. Este último es, sin duda, uno de los aspectos más interesantes de la investigación sobre ED y remite a la importancia de controlar en lo posible las diferencias, quizás inevitables por otra parte, que se producirán en el proceso de mediación en función de los diferentes aplicadores. Otra limitación de este estudio hace referencia al contenido mismo de la prueba dinámica; no es posible construir una tarea de lenguaje «libre de contenido», argumentan los autores. La afirmación anterior introduciría un problema difícil de solventar y sobre el que volveremos más tarde; a saber, que los estudiantes con

alto nivel de vocabulario y más conocimientos previos sobre reglas gramaticales, partirían con ventaja en un proceso de evaluación dinámica de estas características.

Por último, en referencia a la utilidad de la información obtenida con vistas a la intervención, los autores distinguen tres posibles aplicaciones en función de los resultados: (a) alumnado con puntuaciones medias en pretest y bajas puntuaciones en potencial de aprendizaje se beneficiarían de una enseñanza dirigida especialmente a la adquisición de estrategias de aprendizaje; (b) alumnado con puntuaciones medias en pretest y altas en potencial de aprendizaje se beneficiarían de un mayor dinamismo en el material de aprendizaje y mayores oportunidades de aprendizaje autónomo; y, por último, (c) alumnado con bajas puntuaciones en la fase pretest y bajas puntuaciones en la fase de mediación se beneficiarían de trabajar intensivamente aspectos generales de aprendizaje y estrategias de resolución de actividades utilizando materiales con contenidos más básicos de Lengua extranjera para pasar progresivamente a las actividades convencionales de aprendizaje de una segunda lengua.

La Evaluación Dinámica de la aptitud para aprender otro idioma: la prueba CANAL-FT de Sternberg y Grigorenko

La evaluación de la aptitud para el aprendizaje de otro idioma ha tenido lugar tradicionalmente desde enfoques psicométricos. La base de la construcción de las pruebas, destinadas mayormente a la predicción de la capacidad de aprendizaje de los sujetos, ha sido fundamentalmente empírica y, por tanto, no sustentada en la mayoría de los casos en ninguna teoría o conceptualización sobre este tipo de aprendizaje (Sternberg y Grigorenko, 2003). Aunque es cierto que, en general, algunas de estas pruebas han servido de forma eficaz en la predicción del éxito en el aprendizaje de otras lenguas, se han puesto de manifiesto algunos elementos que podrían contribuir a mejorar y generalizar la posibilidad de valorar la capacidad de su aprendizaje. Estos elementos hacen referencia básicamente a: (a) la sustentación de una prueba de evaluación en una teoría cognitiva de la adquisición del conocimiento; (b) la relevancia de la creación de una situación de aprendizaje simulado y gradual de un nuevo idioma en un contexto lingüístico integrado; (c) la relevancia de una comprobación dinámica del aprendizaje del sujeto; (d) la multifuncionalidad de la prueba desarrollada, en el sentido de ser útil, tanto para valorar el nivel de aptitud de los sujetos, como para ofrecer información sobre sus dificultades y puntos fuertes con el fin de diseñar una intervención ajustada; y (e) la posibilidad de realizar nuevas pruebas adaptadas y desarrollar nuevos ítems, por estar basada en la *teoría de respuesta al ítem* (TRI) (Hambleton, 1983).

La concreción de estos postulados tiene lugar en el desarrollo de la prueba dinámica CANAL-FT (*Cognitive Ability for Novelty in Acquisition of Language as applied to Foreign Language Test*) (Grigorenko y Sternberg, 1998), basada a su vez en la teoría CANAL-F, cuyo principal supuesto es que *la adquisición de un idioma extranjero es una forma de experiencia en desarrollo* (Sternberg y Grigorenko, 2003). Esta teoría sostiene también que una de las aptitudes fundamentales para aprender un nuevo idioma es la capacidad para afrontar la novedad y la ambigüedad, lo que a su vez se encuentra recogido en la teoría triárquica de la inteligencia humana formulada por Sternberg.

Desde esta teoría, los procesos de adquisición de conocimiento para el aprendizaje de otro idioma son: (a) *codificación selectiva*, empleada para distinguir la información

relevante, (b) *codificación accidental*, empleada para codificar información secundaria y del contexto base, (c) *comparación selectiva*, que se emplea para determinar la relevancia de la información codificada anteriormente, (d) *transferencia selectiva*, empleada para aplicar reglas y conocimientos aprendidos a nuevos contextos de aprendizaje, y (e) *combinación selectiva*, dedicada a sintetizar e integrar conocimientos adquiridos a través de la codificación selectiva y accidental, así como a la modificación de las estructuras cognitivas que ya posee el estudiante. En definitiva, estos procesos de adquisición de conocimiento se aplican en los cuatro niveles de procesamiento considerados (léxico, morfológico, sintáctico y semántico) y bajo dos modalidades de input y output de la información: visual y verbal.

La prueba dinámica CANAL-FT comprende nueve secciones, cinco de las cuales suponen un recuerdo inmediato, consistiendo el resto en réplicas de secciones anteriores salvo porque son presentadas posteriormente y suponen un recuerdo diferido. Las cinco secciones diferentes de la prueba miden de forma integrada los siguientes aspectos: (a) aprender el significado de neologismos a partir del contexto (recuerdo inmediato y diferido); (b) comprender el significado de pasajes (recuerdo inmediato y diferido); (c) aprendizaje continuo de pares asociados (recuerdo inmediato y diferido); (d) inferencia de oraciones (recuerdo inmediato y diferido); y (e) aprender las reglas del idioma (recuerdo inmediato). Se realizaron dos estudios con esta prueba (Grigorenko, Sternberg y Ehrman, 2000); en el primero de ellos se analizó la relación entre los resultados obtenidos en ésta y en otra prueba de aptitud para el aprendizaje de un idioma extranjero, y también se analizaron las puntuaciones obtenidas en la prueba dinámica en comparación con las otorgadas por una prueba de inteligencia general, con la intención de determinar hasta qué punto la prueba dinámica CANAL-FT podía predecir mejor el aprendizaje de una lengua extranjera que una prueba de aptitud cognitiva general. Por otra parte, también se pretendía conocer qué relación mantenía la puntuación en la prueba dinámica con la experiencia previa en idiomas extranjeros. En el segundo de los estudios se trató de determinar la validez predictiva de la prueba CANAL-FT con respecto al rendimiento en cursos de idiomas extranjeros.

En el primer estudio participaron 158 estudiantes universitarios, a los que se administró un total de cuatro pruebas de aptitud. Estas pruebas fueron las siguientes: una prueba de inteligencia «fluida», el Test de factor «G» de Cattell, nivel 3 (Cattell y Cattell, 1974), una prueba de inteligencia «cristalizada», el *Concept Mastery Test* (CMT; Terman, 1970), una prueba clásica de aptitudes para el aprendizaje de un idioma extranjero, el *Modern Language Aptitude Test* (MLAT; Carroll y Sapon, 1958), y la prueba CANAL-FT. Además se administró a los sujetos participantes un sencillo cuestionario sobre la experiencia previa con idiomas extranjeros. La puntuación de la prueba dinámica se realizó mediante una escala basada en la teoría de respuesta al ítem. Estas escalas permiten afinar la información obtenida sobre cada uno de los participantes y sobre las tareas aplicadas en función de un banco de tareas de dificultad variable que evalúan la misma dimensión.

Las puntuaciones obtenidas en la prueba dinámica se correlacionaron con las del MLAT y con las dos pruebas de inteligencia. Por una parte, se quería conocer la validez convergente de la prueba elaborada, para lo cual se analizaron las correlaciones generales y por escalas con la prueba MLAT; los resultados por escalas mostraron que las correlaciones eran significativamente más altas o, cuanto menos, no eran significativamente más bajas que las obtenidas entre CANAL-FT y las pruebas de

inteligencia. La correlación entre las puntuaciones totales también mostró una validez convergente de la prueba dinámica con el MLAT, al tiempo que una validez discriminadora con respecto a las pruebas de aptitud cognitiva general. Este resultado podría reforzar, bajo nuestro punto de vista, la idea de que las pruebas que recogen aspectos contextualizados y concernientes al dominio específico que evalúan, en este caso, la aptitud para el aprendizaje de un idioma extranjero, predicen mejor el resultado en una prueba relacionada, que aquellas que miden constructos generales como la inteligencia. De los resultados del total de análisis efectuados con respecto a la convergencia y discriminación con las demás pruebas administradas, concluyen los autores que la prueba dinámica es una medida válida de la aptitud para aprender un idioma extranjero, que converge con la prueba de referencia para este tipo de aptitud y que también correlaciona positivamente con las pruebas de inteligencia, aunque sin mostrarse equivalente a éstas.

Con respecto a las relaciones entre las puntuaciones en CANAL-FT y en el cuestionario sobre experiencias previas con idiomas extranjeros, se obtuvo que la cantidad de idiomas que un sujeto conocía y podía hablar y escribir se relacionaba positivamente con la mayor facilidad para la realización de la prueba dinámica, así como con más altas puntuaciones en la misma.

En cuanto al segundo de los estudios presentados se analizó la validez predictiva de la prueba dinámica utilizando como criterio las calificaciones del profesorado. Participaron 63 estudiantes, todos universitarios, a los cuales se administraron las pruebas CANAL-FT y tres secciones del MLAT. En general, se obtuvieron correlaciones significativas entre las puntuaciones de ambas pruebas y las calificaciones otorgadas por los profesores. Los autores concluyen que la prueba dinámica elaborada predice de forma eficaz las valoraciones del profesorado de idiomas extranjeros.

Como posible revisión crítica de estos estudios, que hemos presentado de forma breve, es necesario decir que la consideración de éstos como propuestas de evaluación dinámica obedece en mayor medida a la contemplación del constructo evaluado —la aptitud para el aprendizaje de un idioma extranjero— como una forma de experiencia en desarrollo, y por tanto, como una entidad dinámica y susceptible de ser modificada por la educación; que al diseño y la estructura de la prueba y de la interacción entre evaluador y estudiante en la misma. En realidad, la mayoría de las escalas pertenecientes a la prueba CANAL-FT parecen constituir una sucesión de pruebas estáticas en las que se ofrecen unas instrucciones que orientan la resolución de la tarea, sin que parezca haber una interacción mediadora que facilite una verdadera aproximación a la evaluación del proceso de aprendizaje. Por otra parte, se pretende recoger, aunque solo en función de los resultados obtenidos, información sobre las fortalezas y debilidades del quehacer del estudiante con vistas a una posible intervención centrada en esta información.

En síntesis, estos trabajos recientes, a pesar de que los propios autores confirman la necesidad de una mayor validación con muestras mayores e incluyendo algunos aspectos específicos y generales derivados de las primeras aplicaciones, como medidas de la memoria de trabajo o una mayor gama de idiomas extranjeros que aprender, respectivamente; y a pesar de las posibles limitaciones que pueden presentar por las razones esgrimidas en el párrafo anterior, ofrecen una imagen de la aptitud para el aprendizaje relacionado con el lenguaje que dista de forma relevante de otras

consideraciones más estáticas. Con base en la teoría de Sternberg que considera las aptitudes como formas de experiencia en desarrollo, la aptitud para el lenguaje, al menos en parte, se basaría en la experiencia con formas de procesar la información que pueden desarrollarse y que, por tanto, no constituyen una entidad estática e inamovible. Aunque estas conclusiones, que establecen los autores de los estudios presentados, no han sido demostradas empíricamente, para lo cual harían falta estudios que relacionaran una intervención centrada en la modificabilidad de las aptitudes para el lenguaje con mejoras en la capacidad lingüística, sí reflejan claramente la teoría de base que sustenta estos trabajos sobre evaluación dinámica que, por otro lado, adoptan un enfoque de evaluación de dominios específicos que otorga especial relevancia a la contextualización de los instrumentos de evaluación.

2.4 La evaluación dinámica de la lectura. Algunas experiencias que reabren el camino

Vygotski hizo una distinción particular entre la evaluación en la ZDP de las funciones cognitivas generales y la evaluación en la ZDP basada en los contenidos de las áreas de aprendizaje (Kozulin y Garb, 2001). Mientras que la gran mayoría de las pruebas desarrolladas han dirigido su atención hacia la valoración de las funciones generales, utilizando a menudo tareas similares o extraídas directamente de tests convencionales, el progreso en la ED de dominios específicos en áreas de aprendizaje, ha sido mucho más lento. Una de las razones que podrían estar detrás de este menor interés y desarrollo de esta vertiente de la ED parece ser que la utilización de pruebas estandarizadas ya confeccionadas en una prueba de ED dirigida a dominios específicos, acarrearía diversos problemas que podrían estar relacionados con la calidad «cristalizada» de contenidos específicos como la aritmética o la comprensión de textos y, por tanto, y supuestamente, difícilmente susceptibles de modificación a corto plazo. También adquiere relevancia el hecho de que en el proceso de comprensión tenga una enorme importancia el conocimiento previo del alumno; incluso podría ser a veces superior su relevancia al de otras funciones cognitivas. Todo esto podría explicar, según Kozulin y Garb (2001), que sea necesaria la construcción de materiales específicos que se muestren sensibles a las diversas estrategias que pueden ser evaluadas en la lectura.

El interés concreto por la ED de la lectura es muy reciente y responde fundamentalmente a los cambios en la conceptualización que han ido teniendo lugar con respecto a ésta, así como con relación específica a los procesos de comprensión. La mayor atención a los aspectos *cognitivos* y *metacognitivos* permite un enfoque distinto. Se ha constatado que los «buenos lectores» o lectores estratégicos, ejercerían un control y una regulación constantes sobre el proceso de comprensión, planificando su lectura, supervisando el proceso y evaluando su propia comprensión; ajustando y re-ajustando continuamente objetivos y esfuerzo para conseguir sus metas; el lector pasa a ser un *sujeto activo* que infiere, planifica, razona, integra, supervisa, evalúa... La metacognición va así unida a esta mayor actividad mental constructiva del alumno en el proceso de comprensión. Cada vez más investigaciones, algunas de las cuales revisamos en este trabajo, han relacionado la metacognición con la comprensión lectora, concluyendo, entre otras cosas, que los procesos metacognitivos juegan un papel fundamental en el acceso a una *comprensión crítica e interpretativa* de los textos (Alonso Tapia y Cols., 1992; Repetto, Téllez y Beltrán, 2002; Téllez, 2004). Diversos autores (Palinskar y Brown, 1984; Garner, 1987; Mateos, 1991; Montanero, Blázquez y León, 2002) han mostrado a través de sus estudios que las estrategias metacognitivas se

pueden enseñar y que su instrucción sistemática puede mejorar la comprensión. Un aspecto crucial de la investigación relacionada con la instrucción de estrategias metacognitivas en la lectura, tiene que ver con el efecto de compensación que puede tener la puesta en práctica de estas estrategias sobre la menor cantidad de conocimientos previos que poseen alumnos y alumnas con DA (Dole, Brown y Trathen, 1996).

De los trabajos mencionados anteriormente, pueden desprenderse igualmente unas premisas fundamentales para lograr implementar el uso de estas estrategias en el aula. Estos puntos podrían resumirse básicamente en: (a) dirigir una mayor atención a los procesos de aprendizaje que al producto de la comprensión, (b) proponer tareas o actividades de resolución de problemas, en las que haya que aplicar estrategias, y que éstas hayan sido analizadas previamente por el docente, y (c) proporcionar al alumnado oportunidades para el uso de las estrategias aprendidas en diversas situaciones de aprendizaje, con el objetivo de favorecer la transferencia.

En una reciente propuesta realizada en nuestro contexto, Romero y Lavigne (2006) establecen un procedimiento de evaluación psicopedagógica que incorpora elementos propios de la evaluación dinámica. En este sentido, estos autores proponen la comparación de la información obtenida a raíz de la ejecución individual de tareas por parte de sujetos con dificultades específicas de aprendizaje (por ejemplo, de lectura) con la ejecución en una situación de interacción con el evaluador. Las diferencias en (a) el tiempo de ejecución de la tarea, (b) los errores cometidos, (c) el número de verbalizaciones y peticiones de ayuda, y (d) el número de sugerencias y guías realizadas por el evaluador, constituirán la base para establecer una medida del potencial de aprendizaje, que complementaría el proceso de evaluación de las dificultades del alumno. De este modo, los autores establecen que, en la medida en que el sujeto cometa menos errores en la situación de interacción, emplee un menor tiempo de ejecución, emita menos verbalizaciones y peticiones de ayuda, y necesite un menor número de sugerencias estratégicas para solucionar su tarea, su potencial de aprendizaje sería más alto. Por el contrario, aquellos sujetos que no obtienen beneficios significativos de la mediación del evaluador en la ejecución de la tarea, revelarían un bajo potencial de aprendizaje. Los autores destacan especialmente la validez de la información obtenida de este modo para la planificación instruccional posterior al proceso de evaluación.

Por su parte, Cioffy y Carney (1983) hablan de que las pruebas convencionales de lectura pueden aportar información sobre los conocimientos del alumno y sobre las habilidades de comprensión, pero no pueden otorgar información sobre el potencial de aprendizaje y, sobre todo, suelen ser de poca ayuda a la hora de identificar las condiciones optimizadoras del proceso. En un estudio de caso que presentaron estos autores, se aplicó a un niño de 3º curso de Educación Primaria una prueba dinámica que incluyó en la fase de intervención una fase previa de enseñanza del vocabulario, activación de conocimientos previos, proporción de información relevante sobre el texto, de guías de lectura y de claves textuales para identificar la estructura, así como de claves que permitían entender qué era necesario aprender en la actividad. Los resultados mostraron que la velocidad media de lectura silenciosa fue adecuada y que siete de las ocho cuestiones preguntadas fueron contestadas correctamente. En su lectura en voz alta, el chico cometió solo siete errores, sin que afectaran éstos a aspectos semánticos o sintácticos del texto. La evaluación del potencial de aprendizaje en lectura proporcionó de esta forma, en este caso, información específica para crear condiciones favorables para la mejora del chico.

Otro estudio de caso fue presentado por Kletzien y Bednar (1990), en el que trabajaron con una alumna en edad adolescente que presentaba una desfase significativo en su competencia lectora con respecto a su nivel escolar. Los resultados de una exploración inicial indicaron que la alumna no tenía dificultades relativas a la decodificación del texto y también que toda la información que otorgaba sobre su lectura estaba basada en sus conocimientos previos; cuando estos conocimientos no eran relevantes para dar cuenta de la comprensión, los resultados bajaban notablemente. La ED reveló el uso que esta alumna hacía de la estrategia de *representación visual* o *visualización* en la lectura de textos no escolares, lo que permitió activar el uso de esta estrategia en la lectura de los textos que debía leer en la escuela, anticipando posibles escenas de párrafos posteriores. Se reveló de esta forma el desarrollo potencial de lectura y se ayudó a «ensanchar» la ZDP de la alumna.

Una aplicación práctica con relación al dispositivo de evaluación presentado. La evaluación dinámica de metaconocimientos sobre la comprensión de textos

En un estudio realizado por nosotros (Navarro y Mora, 2003b) nos proponíamos valorar la eficacia y utilidad de la aplicación dinámica de un instrumento de evaluación sobre metaconocimientos en la comprensión de textos con un grupo de alumnos que presentaba dificultades de aprendizaje. El instrumento de evaluación en cuestión, era una de las 32 actividades de las que consta el dispositivo de evaluación dinámica EDPL, que ha sido objeto de aplicación experimental y análisis en la presente investigación. Se trata de una de las tareas pertenecientes al bloque de procesos metacognitivos. En el estudio se buscaba, por una parte, la obtención de información útil para una posterior intervención sobre las DA de la lectura y, por otra, la obtención de información sobre la capacidad mostrada por los sujetos en la resolución de la tarea propuesta. Para tales fines, se pretendía valorar el proceso de resolución de la actividad, los resultados obtenidos y la posible mejora en los aspectos evaluados en la tarea al utilizar esta propuesta. Por otra parte, era necesario también realizar un análisis descriptivo de la cantidad y calidad de la utilización de metaconocimientos por parte de los sujetos evaluados, así como establecer una comparación entre la aplicación grupal del cuestionario EML (Navarro y Mora, 2003a) y su aplicación individual mediante la evaluación dinámica propuesta. La muestra estuvo compuesta por 20 alumnos y alumnas que presentaban DA de la lectura o discapacidad asociada a problemas de comprensión; pertenecían a 5 centros de Educación Primaria y E.S.O. y sus edades estaban comprendidas entre los 9 años y los 16.

Tal y como hemos apuntado, se trataba de una actividad en la que principalmente se pretendían evaluar los metaconocimientos acerca de las estrategias que pueden usarse para comprender los textos, y acerca del propósito y la naturaleza misma de la lectura. Para ello se proponían al alumno una serie de cuestiones de opción múltiple (26 en total) en las que debía elegir, por un lado, la respuesta más acorde con las acciones que pone marcha antes de leer un texto, durante la lectura del mismo, o al finalizar; y, por otro, la respuesta que cree más adecuada para comprender mejor los textos, aunque no coincida con lo que habitualmente pone en práctica. Para facilitar a los participantes una posible respuesta diferenciada con relación a lo que cree más adecuado para comprender, por una parte, y lo que realmente hace cuando se enfrenta a un texto, por otra, el instrumento utilizado contenía dos columnas, en una de las cuales se preguntaba al niño qué estrategias de las que aparecían entre las opciones, ponía habitualmente en práctica. De esta forma, unido a la interacción mediada que ejercería el profesor, en el

sentido de tratar de provocar experiencias de reflexión metacognitiva en el alumno, es decir, experiencias de aprendizaje en las cuales el alumno piense de forma consciente sobre sus propios procesos cognitivos o sobre las estrategias que está utilizando o puede utilizar; se pretendía obtener información sobre las estrategias conocidas y puestas en práctica por éste con relación a la comprensión. El presupuesto de partida era que la evaluación dinámica durante esta experiencia metacognitiva (Mora y Aguilera, 2000), provocada intencionalmente por el profesor y por las demandas de la tarea propuesta por éste, constituye una fuente de información idónea para acceder a los procesos de pensamiento del alumno y detectar dónde se encuentran sus dificultades. Igualmente, introduciendo esta doble posibilidad de respuesta, se pretendía facilitar la oportunidad de la toma de conciencia posterior sobre la diferencia entre lo que se hace y lo que se puede llegar a hacer para comprender mejor. Aunque éramos conscientes de que esto último era ambicioso, nuestra propia experiencia nos confirmaba que durante la fase de mediación o, en su caso, durante la intervención posterior en los problemas detectados, podía profundizarse en esta toma de conciencia por parte del alumno, requisito importantísimo para acceder a mejores cotas de comprensión.

Por otra parte, además de la posibilidad de recoger información sobre los metaconocimientos en la comprensión, la ED de estos procesos a través de la actividad propuesta, nos permitía analizar también los procesos de autorregulación de la propia comprensión; es decir, en qué medida el alumno conoce y utiliza aquello sobre lo que se le estaba evaluando en la prueba de metaconocimientos. La tarea permitía observar y analizar, mediante la interacción adecuada con el alumno, los procesos de planificación, supervisión y autoevaluación de la propia actividad, con lo que el evaluador podría recoger una información más completa y con mayor grado de validez y fiabilidad.

El análisis de las valoraciones sobre la puesta en práctica de los procesos de autorregulación durante la realización de la tarea por parte del alumnado, indicó la presencia de dificultades importantes en el control y regulación del propio proceso de comprensión y, en general, de la propia actividad realizada. La información otorgada por los evaluadores (profesores y orientadores) que aplicaron el instrumento, apuntaba que la propuesta parecía ser útil para obtener información sobre los procesos metacognitivos implicados en la comprensión y para que esta información pudiera ser conducente a la mejora de estos mismos procesos. La mediación efectuada por los evaluadores logró discriminar eficazmente entre los conocimientos sobre aquello que los alumnos creían adecuado para comprender un texto, por una parte, y las acciones que habitualmente ponían éstos en práctica para resolver tareas de lectura, por otra. Así mismo, permitió afinar la valoración que los evaluadores realizaron sobre la presencia o ausencia de una habilidad determinada. De esta forma, aunque la valoración efectuada sobre procesos atencionales de carácter básico fue alta en general, la valoración que se llevó a cabo sobre la puesta en práctica de habilidades de control metacognitivo de estos mismos procesos, mediante la distribución de la atención en función de la tarea o de la relevancia de la información, arrojó resultados muy bajos, matizando y enriqueciendo la información obtenida.

La evaluación realizada, no sólo contribuyó a revelar más información sobre las capacidades y dificultades de los sujetos evaluados con respecto a los contenidos tratados en el cuestionario, así como sobre la puesta en práctica de procesos de autorregulación, otorgando de esta forma la posibilidad de mejorar la planificación y la orientación de una posterior intervención; ocasionó también un incremento en el

rendimiento del alumnado participante, todos ellos con serias DA de la lectura. Más allá de las diferencias significativas encontradas en la comparación efectuada entre el grupo de sujetos evaluados dinámicamente y el grupo de alumnos con DA extraídos de la muestra participante en el estudio en el que se aplicó la misma tarea de forma estática y grupal, este mayor rendimiento con respecto a la realización de la tarea sin la mediación del evaluador, se muestra coherente con los resultados obtenidos en algunos de los estudios sobre evaluación dinámica del potencial de aprendizaje que hemos tenido oportunidad de analizar en este mismo capítulo, como los realizados por Feuerstein y sus colaboradores (Kozulin, 2000). Aunque nuestro estudio no contenía un diseño con pretest y postest en el que poder certificar cuantitativamente posibles mejoras en la actuación de los sujetos, la información otorgada por los evaluadores, relativa a las mejoras observadas en el rendimiento sobre los contenidos evaluados cuando éstos fueron valorados con la propuesta de evaluación dinámica, apunta claramente a la existencia de un incremento significativo y a que éste sería deudor en gran medida de la metodología empleada.

En este sentido, el análisis efectuado sobre la propia puesta en práctica de la evaluación dinámica de metaconocimientos sobre la comprensión, en sujetos con DA de la lectura, reveló la obtención, a juicio de los evaluadores, de una información cualitativa de mejor calidad que la que pudo obtenerse con la pasación de una prueba similar en forma estática. Los evaluadores destacaron cuestiones relativas, por ejemplo, a procesos de ajuste personal-social en los alumnos: afrontamiento de la tarea, estados de motivación, sensación de incompetencia, rechazo manifiesto y sincero por las tareas de lectura... Así mismo informaron sobre la puesta en práctica de procesos cognitivos y procesos de control y regulación de la propia actividad lectora. En general, la información obtenida fue considerada útil por los evaluadores para una posterior intervención en las DA del alumnado; y el método propuesto fue considerado en general eficaz para obtener esta información más precisa y contextualizada, que permitía abordar la solución o mitigación de los problemas.

3. La evaluación dinámica de procesos implicados en la lectura. Bases teóricas y metodológicas de nuestra propuesta y objetivos de la investigación

3.1 Bases teóricas y metodológicas de nuestra propuesta

Bases teóricas y metodológicas generales de la evaluación dinámica de la lectura: Cuestiones teórico-prácticas y psicométricas de base

Parece ser que el objetivo primordial que ha presidido la mayoría de las investigaciones en el ámbito de la ED es el de demostrar que el pensamiento en acción, o mejor, en interacción, guiado o dirigido hacia el aprendizaje y mediante la ayuda ajustada, es el modo más adecuado de evaluar el desarrollo potencial del aprendizaje. Más allá de este objetivo general, que podríamos calificar como consustancial al enfoque dinámico, uno de los objetivos más específicos que se han propuesto muchos de los modelos de ED ha sido el de ofrecer una mejor valoración del constructo de una determinada aptitud (Embretson, 1987). En nuestro caso, nos proponemos contribuir a optimizar la evaluación de los procesos implicados en la lectura; para esto ha sido necesario un análisis exhaustivo de los procesos y elementos implicados en ésta y sobre cómo podían ejercer influencia en la valoración que hacíamos de un alumno, desde procesos relativos a las funciones cognitivas del sujeto, a otros de tipo contextual relativos al aula o la familia, pasando por aspectos relacionados con la motivación o las actitudes.

El análisis de estos elementos no conducirá habitualmente a un control exhaustivo de los mismos en el proceso de evaluación en el contexto escolar, pero sí puede proporcionar información muy valiosa a la hora de interpretar los resultados obtenidos. Según Sternberg y Grigorenko (2003), entre los requisitos de las pruebas dinámicas para responder a una valoración del proceso de aprendizaje que mejore significativamente la ofrecida desde las pruebas estáticas, estaría el de encajar bien con un «modelo de atributos latentes», es decir, tener en cuenta durante la evaluación y en la interpretación de los resultados la existencia de factores socio-culturales y socio-emocionales que cuanto menos condicionan la situación de prueba. Otro requisito destacado por estos autores es el de la validez predictiva de las pruebas. Ya apuntábamos anteriormente, cuando hablábamos de las diferentes pruebas y los diferentes modelos dinámicos revisados, que la validez de predicción, especialmente sobre el rendimiento escolar o el rendimiento en situaciones de aprendizaje prolongadas, ha sido uno de los aspectos que más ha preocupado a los investigadores; sobre este asunto, las medidas criterio adoptadas para valorar la relación entre la valoración del proceso de aprendizaje en una prueba dinámica y el rendimiento, han sido variadas, entre ellas la correlación entre las puntuaciones dinámicas (compuestas por las puntuaciones obtenidas en el conjunto de tareas propuestas, por las puntuaciones posttest, o por la diferencia pretest-posttest, básicamente) y las puntuaciones obtenidas en pruebas de rendimiento, programas de intervención, o a través de las calificaciones escolares; y, por otra parte, las correlaciones entre puntuaciones dinámicas y la valoración del profesorado acerca de la

competencia del alumno en la aptitud de que se trate. En nuestro caso, hemos optado por valorar el nivel de predicción del dispositivo de evaluación dinámica de procesos implicados en la lectura (EDPL) sobre el rendimiento escolar, mediante las calificaciones escolares obtenidas en el área de Lengua al finalizar el curso, así como mediante la valoración del profesorado sobre el rendimiento, la competencia y la actitud en tareas de lectura. Un tercer requisito destacado por Sternberg y Grigorenko es la necesidad de tener una idea clara de cómo influyen los factores relativos a la interacción de procesos y estrategias, mediados por otros factores socio-culturales y socio-emocionales, no sólo en la situación de prueba, sino en la actuación del alumno.

Por otra parte, otro objetivo específico que se refleja explícita o implícitamente en muchas pruebas dinámicas es la mejora de las funciones cognitivas de los sujetos. Es en este sentido en el que la ED se aleja significativamente de los objetivos que presiden en gran medida las pruebas convencionales; la ED conecta evaluación e intervención probablemente no porque constituyera una meta inicial de estos procedimientos, sino porque precisamente en el *proceso de intervención* es donde se considera más valioso el análisis de las funciones cognitivas. En este caso, las pruebas evalúan la modificabilidad y el cambio, en qué sentido se produce, cuál es la calidad y la cantidad de ese cambio, cómo tiene lugar y mediante qué tipo de intervención; todo esto con el objetivo de orientar hacia la acción educativa óptima para seguir desarrollando el potencial de aprendizaje. Esta opción de la ED no es la única, aunque sí nos parece la más interesante. De otro modo, como expresa Saldaña (2001), un modelo dinámico interesado exclusivamente en medir y analizar las causas del cambio producido entre una fase pretest y una postest, ofreciendo quizás un índice cuantitativo de este cambio, corre el riesgo de ser interpretado como una nueva medida estática de un nuevo constructo (modificabilidad, ZDP, potencial de aprendizaje...) con pretensiones de sustituir a otro (inteligencia).

Lógicamente, debe tener lugar la investigación que valore y analice el cambio y la modificabilidad de un sujeto en una situación de aprendizaje; la cuestión es con qué objetivos: el de sintetizar una nueva medida estática de la inteligencia o similar, o el de proporcionar información valiosa para la mejora del proceso de enseñanza y aprendizaje. De acuerdo con Saldaña y como hemos expresado ya en otro lugar (Navarro, 2004), lo relevante y específico de la ED consistiría en evaluar las funciones cognitivas o socio-personales del alumno, situados en su ZDP, ajustando nuestra intervención para obtener información sobre los procesos que el alumno utiliza en la resolución de las actividades propuestas, observando y analizando cómo hace uso de estos procesos y determinando cuál es la mejor manera de ayudarle a que mejore su actuación. Si pretendemos que la investigación relacionada con la ED, en consonancia con la intención y el trabajo inicialmente desarrollado por Vygotski y posteriormente continuado por la escuela soviética, sirva para la mejora de los procesos educativos, debemos situarnos claramente en esta opción. Algunas dificultades pueden, sin embargo, aparecer cuando la ED está enfocada a esta mejora del proceso de enseñanza y aprendizaje al que aludíamos. Estas dificultades, bajo nuestro punto de vista, deben ser conocidas —y en la medida de lo posible, minimizadas— y asumidas como consustanciales al hecho educativo y a los procesos de enseñanza y aprendizaje en sí mismos cuando estos son considerados en toda su complejidad y dinamismo. Por ejemplo, en educación, el orientador o el profesor de apoyo encontrarán algunos problemas para poder implementar pruebas dinámicas de forma individualizada y siguiendo unos métodos excesivamente rígidos o estructurados; el objetivo en estos

casos es evaluar para la mejora, y puede subordinarse la comprobación y la obtención exclusiva de medidas a la modificación y la determinación de soluciones. Aún en el curso de una investigación como la nuestra, donde los participantes (orientadores y profesorado de apoyo a la integración) conocían la relevancia y necesidad del control de determinadas variables, hemos constatado algunas de estas dificultades que expresan sin más la realidad educativa en la que deben ser puestas en práctica las propuestas construidas precisamente con el objetivo de mejorarla. A pesar de las dificultades expuestas, parece ser que esta opción decidida por la intervención y la mejora de los procesos educativos está ganando terreno entre las investigaciones y propuestas de los últimos años. Esto se observa en el cambio que se ha ido produciendo desde pruebas o estudios orientados exclusivamente a la predicción o la medición del potencial de aprendizaje, hacia pruebas y propuestas metodológicas más orientadas al aprendizaje (Delclos et al., 1992; Ruijsseenaars, Castelijns y Hamers, 1993; Sternberg y Grigorenko, 2003).

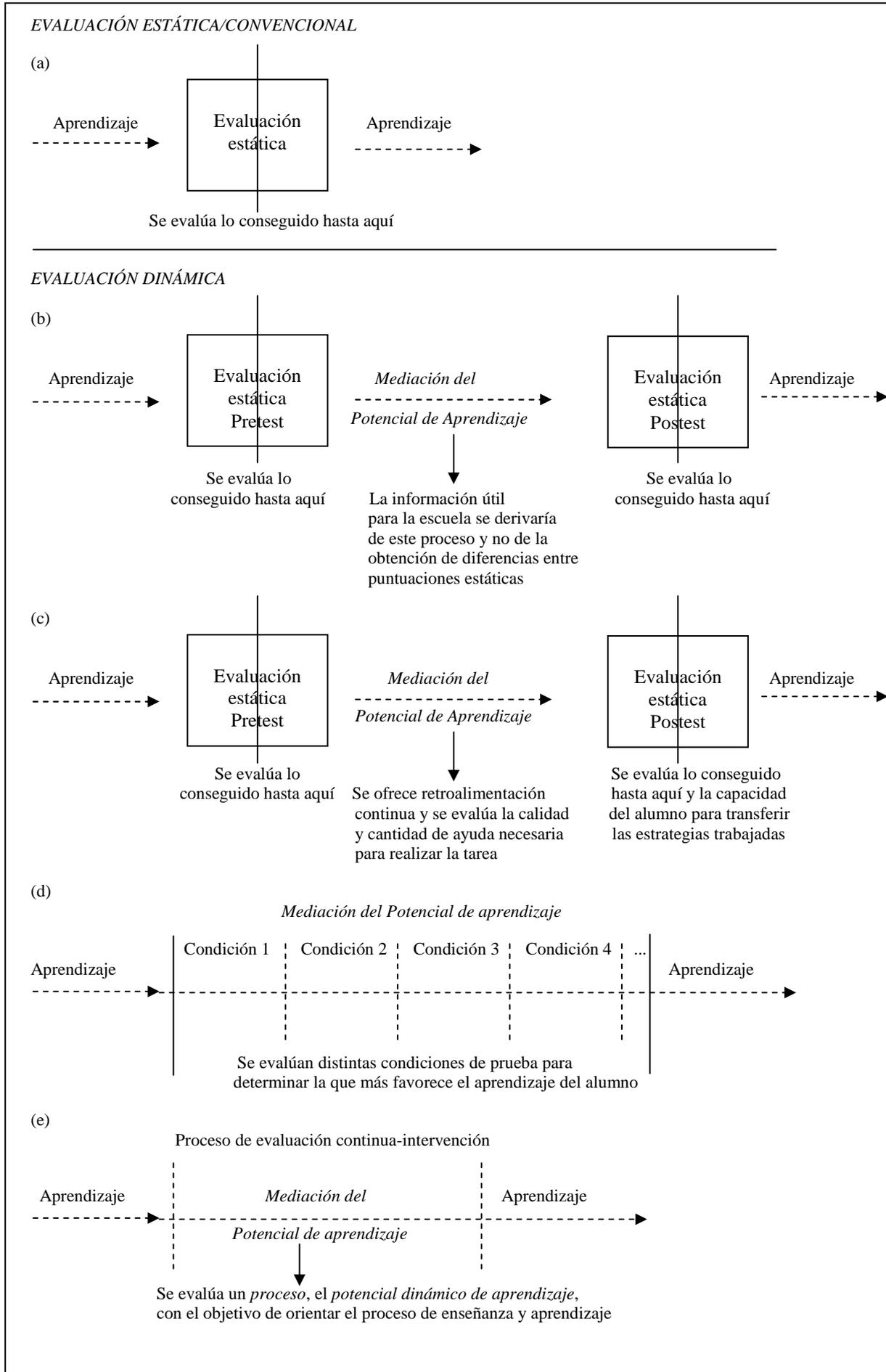
Además de las dificultades prácticas que puede conllevar la adopción de enfoques dinámicos, dificultades que se derivan de una concepción de los procesos educativos más compleja y contextualizada, otro de los aspectos que más revisiones críticas ha suscitado es el que tiene que ver con cuestiones relativas a la fiabilidad y validez de las medidas criterio para establecer las mejoras obtenidas por un sujeto, en otras palabras, con la medición del cambio. Ya hemos comentado el mayor interés que progresivamente ha ido tomando este asunto entre los investigadores de la ED. El contenido «estándar» de muchas pruebas dinámicas, como hemos tenido oportunidad de ver, se basa de forma aproximada en la aplicación de tests de inteligencia convencionales, seguidos de un entrenamiento en las estrategias o habilidades consideradas idóneas para superar los ítems de los tests, realizado con tareas similares, y, finalmente, un postest compuesto por las mismas pruebas iniciales o pruebas paralelas. Este prototipo que, a nuestro parecer, presenta ciertas limitaciones, dio pie para que se analizaran algunos de los factores relevantes que podían explicar el cambio. En este sentido, algunos estudios han mostrado que la simple repetición de la prueba inicial explicaba ya de forma significativa las mejoras en el postest para el 30% de los niños (LeGagnoux et al., 1990; Klauer, 1993); esto indicaría que el componente o efecto de práctica explicaría una parte importante de los resultados. En este sentido, entre las precauciones que minimizarían este efecto de la práctica se encuentra el contar con un grupo de control que permita la comparación de los cambios producidos por la intervención y, al mismo tiempo, de todos los efectos que pueden acompañar a la aplicación de la prueba. Otra forma de evitar estos problemas reside en la propia fase de construcción de la prueba; de este modo, los dispositivos dinámicos que no se basan exclusivamente en modelos *pretest/ mediación/ postest* no presentarían esta limitación. También relativo a la construcción de la prueba, parece relevante que el contenido de las actividades de evaluación dinámica no debería guardar una relación de semejanza tan acusada con el pretest y el postest; en estos casos, estas pruebas serían las que presentarían mayores problemas para ejercer un control efectivo sobre el efecto de la práctica. Por otra parte, parece lógico pensar que en un procedimiento de evaluación-intervención que contiene un entrenamiento estratégico que se aplica con el fin de mejorar la actuación en una determinada aptitud; mientras más distancia haya entre las actividades de la fase de intervención y las tareas criterio de las pruebas postest, mayor será la ganancia obtenida; expresada, no sólo en términos de diferencias pretest-postest, sino también en términos de *transferencia*.

Sternberg y Grigorenko (2003) hacen alusión a los problemas que puede representar la medición del cambio, incluso en el caso de que estas potenciales dificultades de tipo metodológico estuvieran controladas; en efecto, se ha destacado por parte de los expertos en psicometría de orientación dinámica, el hecho de que la mayoría de los supuestos que deben cumplir las pruebas psicológicas para admitir modelos estadísticos clásicos, como el modelo de regresión, no se cumplen en las pruebas dinámicas comúnmente utilizadas (Schöttke, Bartram y Wiedl, 1993; Sijtsma, 1993; citados en Sternberg y Grigorenko, 2003). En este sentido, quizás no sólo estos modelos estadísticos clásicos se muestren ineficaces para medir el cambio producido en el curso de un procedimiento de evaluación dinámica, sino que tal vez tuviéramos que extender esta reflexión hasta lo que puede suponer ciertamente un contrasentido, esto es, el hecho de que se utilicen pruebas de carácter estático, que evalúan aptitudes mayormente ya cristalizadas y de forma individual, para valorar el desarrollo potencial del aprendizaje, lo que supone ciertamente evaluar un proceso. A nuestro juicio, esta es la mayor limitación de la utilización de modelos clásicos de evaluación para la medición del cambio en los enfoques dinámicos. Si bien es cierto que el resultado de un postest en el que se registre una mejora de la aptitud valorada otorga información sobre una incorporación o integración de la información y/o sobre los procesos entrenados que han debido ser puestos en práctica en su ejecución, no nos ofrece, en cambio, información sobre lo que parece tener mayor relevancia en la ED, aquello para lo que han sido diseñados y contruidos los dispositivos de ED, es decir, sobre la propensión o la capacidad de un sujeto para beneficiarse de la mediación ejercida por un adulto o compañero aventajado con vistas a mejorar su actuación; en otras palabras, no nos informa sobre su potencial —siempre dinámico— de aprendizaje. En la figura 2 podemos observar una comparación entre distintos enfoques de evaluación.

En efecto, si la acción educativa tiene lugar en contextos de interacción, la evaluación debe recoger y valorar sobre todo la actuación del sujeto en un contexto y una situación de interacción; por tanto, en los procesos de ED debería ser primordial la utilización de procedimientos dinámicos para evaluar el cambio producido en el sujeto, con el objetivo principal de orientar acerca de la optimización de la propia acción educativa. En este sentido, la ED, dentro del proceso de enseñanza y aprendizaje, actuaría como un corte que realizamos —con un espacio y tiempo determinados de duración— en un proceso de evaluación continua e intervención (apartado (e) de la figura), con el fin de obtener una información valiosa dirigida a orientar y guiar el proceso, o con el fin específico de la elaboración de un informe que oriente la modalidad de escolarización de un alumno o los estudios en los que su potencial dinámico de aprendizaje se vería más desarrollado a priori.

Hemos hablado, a raíz de la revisión de algunos modelos dinámicos, de la relevancia que la validez ecológica de los resultados ha tomado dentro de la psicología educativa. La validez ecológica de una prueba podríamos determinarla en función de los contenidos de la misma, es decir: si una prueba evalúa contenidos que son trabajados en la escuela, podríamos decir que los resultados poseen validez ecológica porque se ha tenido en cuenta el contexto habitual de actuación para el diseño de la prueba en cuanto a su contenido, y la medición de los resultados obtenidos informará sobre la capacidad que los sujetos han desarrollado sobre estos contenidos. Por otra parte, también podríamos determinar la validez ecológica en función del procedimiento de evaluación; en este sentido, en la medida en que el procedimiento utilizado se asemeje a los procedimientos de enseñanza y aprendizaje trabajados en la escuela, podríamos

Figura 2. Comparación gráfica entre distintos modelos de evaluación del aprendizaje



hablar de una mayor validez ecológica del proceso y de los resultados. Es especialmente en este segundo punto donde la ED se destaca sobre los procedimientos estáticos de evaluación; sin embargo, para que esto demuestre ser cierto es necesario, por un lado, conocer los procedimientos de enseñanza y aprendizaje escolar y también los contenidos trabajados, y diseñar de forma contextualizada las pruebas de evaluación, y, por otro, proceder a una evaluación del proceso de aprendizaje, porque es justamente esto, aprender, lo que constituye el eje fundamental de la acción que desarrolla el alumnado en la escuela. De todo esto se deriva que los estudios encaminados a valorar el poder predictivo de las pruebas dinámicas pueden haberse encontrado con serios problemas cuando se intenta correlacionar las medidas del potencial de aprendizaje obtenidas en situaciones de interacción colaborativa y dirigidas en ocasiones a procesos y habilidades cognitivas y metacognitivas, con medidas del rendimiento escolar obtenidas quizás con procedimientos de enseñanza cuanto menos no conocidos, o incluso con medidas estáticas del rendimiento escolar. Esto al menos introduce ciertas dudas acerca de la validez de las conclusiones sobre la validez predictiva de las pruebas obtenida de esta forma. En cambio, como hemos podido comprobar, cuando la predicción se ha realizado sobre situaciones de aprendizaje prolongado que guardaban relación conceptual y metodológica con los contenidos y los procedimientos de la ED, los resultados parecen ser más consistentes y prometedores (Jensen, 2000; Lidz, 2000; Saldaña, 2001; Sternberg y Grigorenko, 2003). En otras palabras, cuanto mayor es la relación entre el contenido de ambos procesos, evaluación e intervención educativa, mayor es el poder de predicción de la prueba dinámica.

Otro aspecto que nos interesa especialmente con respecto al tema de la validez ecológica y la capacidad de predicción de las pruebas dinámicas, es la especificidad de sus contenidos y la contextualización de los mismos. En este sentido, investigadores como Ruijsenaars et. al. (1993) o Hamers, Pennings y Guthke (1994) han desarrollado tareas dinámicas que evalúan contenidos específicos como la decodificación fonológica. Como sabemos, el mayor desempeño en estas tareas relacionadas con la conciencia fonológica está asociado con mayores aptitudes y logros en la comprensión lectora (Jiménez y Ortiz, 1998). Los resultados del estudio de Hamers y sus colaboradores revelaron que la prueba dinámica específica diseñada predecía mejor el rendimiento escolar que otras pruebas generales del potencial de aprendizaje y que otras pruebas de carácter estático. En esta línea, nuestro dispositivo de Evaluación dinámica EDPL incorpora diversas actividades de evaluación que han sido desarrolladas para distintos ámbitos del aprendizaje de la lectura. A nuestro juicio, la inclusión de estas actividades específicas de aprendizaje —actividades que habitualmente son trabajadas y valoradas en los procesos de enseñanza y aprendizaje escolar —en procedimientos de ED dirigidos a contenidos específicos, es especialmente relevante para la obtención de información ecológicamente válida.

De acuerdo con Schlee (1985), citado por Sternberg y Grigorenko (2003), la validez ecológica debe ser el parámetro fundamental a la hora de implementar cualquier procedimiento de evaluación en el ámbito escolar. Esto lleva a este autor a *fusionar* los procesos dinámicos y estáticos de evaluación, destacando una aplicación de éstos orientada al tratamiento y del cual se derive una intervención educativa adecuada y ajustada al contexto escolar. No compartimos, sin embargo, la objeción que presentan Sternberg y Grigorenko en el sentido de que una evaluación orientada a la intervención se convertiría en parte integrante de la didáctica y por tanto carecería de relevancia la cuestión de la validez ecológica; por el contrario, pensamos que la evaluación dinámica,

en el marco de los procesos de enseñanza y aprendizaje, como concreción de un proceso de evaluación continua, tiene sentido en la medida en que puede orientar la intervención educativa, ofreciendo información útil sobre el desarrollo de las funciones cognitivas del sujeto, sobre su potencial dinámico de aprendizaje, así como sobre aquellas pautas metodológicas de mediación que optimizan el rendimiento. Lógicamente, incluso una evaluación que tiene lugar en el marco de los procesos educativos y está orientada al tratamiento, puede otorgar información que trascienda éste y sea útil, en el marco de diseños de investigación, para mejorar los propios procesos de evaluación o para incrementar nuestro conocimiento sobre el funcionamiento cognitivo. En la ED se intenta evaluar un proceso activo de aprendizaje; por tanto, a medida que las condiciones y la situación de prueba otorguen más y mejores posibilidades de actividad mental constructiva de los sujetos, más idónea será la situación para la evaluación del desarrollo potencial dinámico del aprendizaje (Navarro, 2004) y, al mismo tiempo, mayor poder predictivo tendrán las medidas obtenidas con respecto a otras situaciones de aprendizaje.

Nuestro objetivo no es solo predecir, sino evaluar para la mejora del proceso educativo. También podemos contribuir a la introducción de cambios en los sistemas educativos si observamos reiteradamente resultados de la investigación que nos hacen dudar sobre las prácticas de evaluación o instrucción utilizadas en general o con algún grupo de alumnos en particular. La predicción de la actuación o el rendimiento del alumnado tendrá sentido únicamente en la medida en que contribuya a validar un procedimiento de evaluación y, con esto, la información que de él puede desprenderse. El objetivo, como hemos planteado, debe dirigirse más a evaluar para orientar hacia la optimización del desarrollo potencial dinámico del aprendizaje.

La evaluación dinámica, por tanto, es un proceso de evaluación del potencial dinámico del aprendizaje de un sujeto con el objetivo principal de orientar la acción educativa hacia la optimización del proceso de enseñanza y aprendizaje. Los puntos esenciales que a nuestro juicio quedan recogidos en esta definición o que de algún modo la sustentan, son estos:

- La ED es un proceso de evaluación del potencial dinámico del aprendizaje;
- se realiza situados en la ZDP del sujeto, porque es ahí donde puede observarse un proceso activo de aprendizaje;
- se recoge información sobre la capacidad de aprendizaje y sobre la propensión de un sujeto a beneficiarse de la guía y ayuda de un adulto o compañero aventajado;
- se recoge información sobre los procesos cognitivos y/o socio-personales que un sujeto es capaz de desarrollar de forma autónoma y también sobre aquellos que puede desarrollar en colaboración;
- se recoge información sobre los patrones de mediación/intervención más idóneos para favorecer el aprendizaje, así como sobre las condiciones relativas al contexto que optimizarían el proceso educativo;
- el potencial de aprendizaje es un concepto dinámico que se desplaza en función de la adquisición y desarrollo de nuevos aprendizajes;
- la información recogida debe ser valiosa para orientar la acción educativa hacia la optimización del proceso de enseñanza y aprendizaje.

Al comienzo del capítulo dedicado a la Evaluación Dinámica, hacíamos referencia a la *intencionalidad* como uno de los parámetros básicos de la *Experiencia de Aprendizaje Mediado* (EAM) de Feuerstein. Esta intencionalidad era una característica del patrón de interacción mediada que mantenía el adulto con el niño, manipulando intencionalmente los estímulos y las situaciones para provocar la EAM. Este elemento se encuentra presente en nuestra propuesta de ED de la lectura, centrando nuestra atención y la del sujeto, más allá de la actividad concreta que se esté realizando, en los procesos de pensamiento que se ponen en juego; es más, nuestra atención se ha dirigido especialmente hacia los procesos metacognitivos en el transcurso de cada una de las actividades. Los evaluadores tenían orientaciones metodológicas sobre el proceso de mediación y sobre la valoración de procesos metacognitivos que le conducían a focalizar sus observaciones y su mediación hacia estos procesos. Por otra parte, en línea con lo argumentado a lo largo de este trabajo, la capacidad de *apropiación* de estos procesos o, en otras palabras, la capacidad de modificabilidad que muestre el sujeto en el sentido de incorporar e interiorizar determinados procedimientos en su repertorio estratégico que incrementen sus posibilidades de acceso a la comprensión de los textos, se constituye en un elemento revelador del funcionamiento cognitivo actual y también de su capacidad de beneficiarse de aprendizajes futuros.

Cuando planteábamos la *trascendencia* como otra de las cualidades descritas por Feuerstein (1990) en su teoría de la EAM, hacíamos alusión a la relevancia que podía tener la regulación y el control consciente en el planteamiento y la organización de las actividades de mediación. Decíamos también que en función de características como los contenidos del aprendizaje y los sujetos, esta característica de la consciencia iba a ser de suma importancia; en efecto, aunque compartimos con Feuerstein la idea de que determinados contenidos culturales pueden ser transmitidos a través de la participación en actividades sociales (Feuerstein, 1990; Kozulin, 2000) sin que exista un control consciente y deliberado por parte de los miembros expertos; sin embargo, el problema que se nos plantea es que la ausencia de control y regulación consciente de la actividad mediadora no permitiría, obviamente, el control y la regulación del proceso ni, así mismo, su evaluación. Cuando hablamos de educación escolar o, más concretamente, de la especificidad del proceso de enseñanza y aprendizaje escolar (Ausubel, Novak y Hanesian, 1978; Coll, 2001), y aún siendo conscientes de la existencia y la relevancia del llamado *currículum oculto*, parece tener una enorme importancia el hecho de que el profesorado: (a) participe activamente en la planificación de su acción educativa en función de unos objetivos educativos determinados; (b) mantenga, durante el desarrollo de la interacción educativa un control y regulación consciente del proceso, pudiendo de este modo avanzar en el desarrollo de su planificación establecida o introducir modificaciones en función de los acontecimientos; (c) pudiendo, así mismo, valorar durante y al final del proceso, el grado de adecuación del mismo en función de los objetivos planteados. Esto cobra aún mayor importancia cuando de lo que hablamos es de enseñar contenidos relacionados con el aprendizaje de estrategias metacognitivas a alumnado que presenta DA. Hemos hecho alusión durante este trabajo y también en otros (Navarro y Mora, 2002; Navarro, 2004) a la necesidad de una intervención marcadamente intencional en la enseñanza de habilidades metacognitivas a alumnado con DA (Vygotski, 1931/1997; Campione, Brown y Ferrara, 1987; Martín y Marchesi, 1990; Mora, 1998); bajo nuestro punto de vista, para que sea eficaz, este tipo de intervención solo puede realizarse desde planteamientos educativos que, a su vez, asuman estrategias metodológicas de planificación, control y regulación conscientes del proceso, así como la reflexión sobre el propio proceso de enseñanza y aprendizaje como

la manera idónea de introducir mejoras. Por tanto, tal y como planteaba el propio Vygotski (1979/1996), la interiorización de aprendizajes relacionados con las funciones de autocontrol y autorregulación tendría que ver con la forma en que el docente desempeña esas mismas funciones en el curso de su labor. En nuestra propuesta, el valor de esta cualidad de trascendencia aplicada a la mediación de los procesos implicados en la lectura y especialmente de los procesos metacognitivos, se pone de manifiesto cuando centramos nuestra atención en procesos cognitivos, metacognitivos y relacionados con la motivación, la actitud y los intereses, más allá de la actividad concreta que se está desarrollando, y que, en la mayoría de los casos, sirve de soporte para el trabajo sobre los procesos mencionados. En realidad, la asunción de la necesidad o la conveniencia de plantear la interacción mediada de forma consciente y deliberada, conecta estrechamente el valor de la trascendencia con el de la intencionalidad. En suma, se trataría de trascender la actividad de lectura y sus contenidos para ir más allá, facilitando la apropiación de estrategias generales y específicas de aprendizaje; y trabajando igualmente sobre aspectos como las actitudes, las expectativas o el autoconcepto.

Por otra parte, lo esencial de nuestra propuesta consiste en mediar la adquisición de conocimiento metacognitivo y la potenciación del uso de estrategias metacognitivas relacionadas con los procesos de lectura y comprensión; para esto deben establecerse determinadas condiciones (Novak, 1998) que facilitarían y orientarían estos aprendizajes en el sentido propuesto ya por Feuerstein (p.e. 1991) en su teoría de la EAM, es decir: la selección, estructuración y secuenciación de los contenidos a aprender. Por otra parte, en consonancia con las ideas y propuestas que se derivan de las actualizaciones de la teoría del *Aprendizaje verbal significativo*, la mediación que se establece va orientada, entre otros aspectos, a la activación de conocimientos y experiencias previas y a la facilitación de la integración del nuevo contenido de aprendizaje con los conocimientos que ya se poseen, otorgando oportunidades para analizar y revisar lo que ya se sabe y lo que se aprende desde lo que ya se conoce; estableciendo también actividad mediadora encaminada a favorecer una actitud proclive al aprendizaje. Lógicamente, la relevancia de establecer patrones de mediación sensibles a los elementos aportados por la teoría del Aprendizaje verbal significativo se torna de enorme importancia en función de lo fundamental que resultan estas aportaciones en la explicación de los procesos de enseñanza y aprendizaje escolar en la actualidad.

Aprendizaje y evaluación de la lectura: La relevancia del contexto social e histórico

Hace ya más de tres décadas, en sus estudios transculturales, Cole y sus colaboradores (Cole, Gay, Glick y Sharp, 1971; Cole y Scribner, 1974) situaban las diferencias individuales en las diversas formas en que los contextos socio-culturales de referencia promovían o facilitaban la apropiación de sus contenidos relacionados con la cultura. Esto le llevó posteriormente a concluir que las consecuencias de la escolarización y del empleo de instrumentos simbólicos, asociados a los procesos de educación formal —como por ejemplo la lectura o la escritura—, para el desarrollo cognitivo de los niños y niñas, no tenían un carácter absoluto. La estructura y la relevancia de las actividades sociales en una cultura determinada influirán notablemente en el papel que desarrolle un contenido de aprendizaje; así la lectura, por ejemplo, ha desempeñado históricamente un rol muy distinto en el proceso educativo aún dentro de una misma cultura de referencia y, más aún, desempeña distintos roles en el mismo momento histórico pero en diversos contextos socio-culturales. Las aptitudes cognitivas

relacionadas con la escolarización, como recoge Kozulin (2000), y los instrumentos psicológicos desarrollados históricamente y culturalmente, que permiten y facilitan la apropiación de elementos culturales, han adquirido un mayor peso social cuando la estructura de la sociedad ha generado mayores cotas de actividad relacionada precisamente con los contenidos de tipo escolar o académico (Cole, 1990).

En este sentido, la gran mayoría de actividades desarrolladas actualmente en los centros educativos se vertebran en torno al lenguaje escrito. El aprendizaje y dominio de la lectura y la escritura ha ido cobrando progresivamente una mayor relevancia dentro del marco de actividades escolares, debido fundamentalmente a lo esencial que resulta su óptimo desarrollo para garantizar el acceso a los distintos contenidos correspondientes a las diversas áreas de conocimiento. En consonancia con lo que puede observarse en el contexto escolar, Sánchez (1999) hace referencia a que las demandas acerca de lo que es necesario aprender con relación al lenguaje escrito y a través de él, ha ido paulatinamente incrementándose en las últimas décadas. La importancia del dominio de la lectura y la escritura es tal en la mayoría de las sociedades actuales, que la inclusión y la adaptación social de los individuos puede estar influida o condicionada notablemente por este factor (Tolchinsky, 1990).

Unido a esta creciente demanda social a la que hacíamos mención anteriormente sobre la adquisición de la lectura y la escritura, se encuentra un hecho, más bien un proceso, de la reciente historia de la psicología educativa. El aprendizaje de la lectura y su estudio por parte de la psicología ha experimentado en los últimos treinta años un cambio significativo. Con la aparición en escena de la psicología cognitiva, el redescubrimiento de la obra de Vygotski, y algunas aportaciones que evitaban olvidar componentes emocionales; el estudio de la lectura y, por tanto, el estudio de su aprendizaje, ha ido cobrando una relevancia notable en el panorama científico de la psicología relacionada con los procesos educativos. Uno de los aspectos que más trabajos ha concentrado ha sido el de la comprensión lectora. Por una parte, se ha ido enriqueciendo la conceptualización del proceso de comprensión, incluyendo en su desarrollo y adquisición factores de tipo contextual, emocional, socio-cultural, lingüístico, cognitivo y metacognitivo; e igualmente, las dificultades de aprendizaje ligadas a la comprensión, han sido consecuentemente abordadas de manera más acorde con la complejidad de este proceso. Paulatinamente, han ido apareciendo programas educativos para la mejora de la comprensión lectora y, progresivamente, algunos de ellos están llegando a las aulas, aunque su puesta en práctica esté siendo, en la mayoría de las ocasiones, poco continuada y sistemática. Debido en parte a este hecho y a la dificultad de evaluar programas de esta índole, existen aún pocas investigaciones que puedan ofrecer datos más concluyentes sobre la eficacia de estos programas. No obstante, podemos encontrar algunos estudios que analizan elementos pertenecientes a estos programas de comprensión, ofreciendo datos sobre la repercusión de su implementación en la ejecución posterior de los sujetos experimentales. Una revisión de estos estudios se recoge en el volumen de anexos de este trabajo (Anexo XV).

Algunos de estos programas se han desarrollado con la intención de favorecer la adquisición de estrategias de aprendizaje por parte del alumnado; estrategias que les permitan saber cómo identificar la estructura textual, comprender la intención del autor, poner en relación con los contenidos del texto sus experiencias y conocimientos previos, revisar la información obtenida a raíz de sus propios conocimientos, verificar la información, identificar ideas relevantes, formarse opiniones y ser capaz de valorar

críticamente la información, controlar y regular el propio proceso de comprensión, etc. En definitiva, estos programas buscan desarrollar la capacidad de aprender a través del texto, de leer para aprender, y, en ocasiones se busca relacionar el trabajo sobre la lectura con el objetivo más amplio de aprender a aprender y con la enseñanza del pensamiento. (Alonso Tapia, 1992; Isabel Solé, 1992; Maruny, Ministral y Miralles, 2001; Repetto, 1997; Hernández y Quintero, 2001; Vidal-Abarca y Gilabert, 1991; Repetto, Téllez y Beltrán, 2002; Repetto, Sutil, Manzano, Téllez y Beltrán, 2002).

Evaluación de procesos: algunos problemas y propuestas de innovación en la evaluación de la comprensión

En consonancia con el auge que en las últimas décadas ha tenido lugar respecto de los trabajos encaminados a favorecer la adquisición y el desarrollo de la comprensión lectora, fruto de renovadas concepciones sobre la lectura y sobre el proceso mismo de comprensión, aunque en menor cantidad, ha tenido lugar el desarrollo de pruebas de evaluación. Estas pruebas no siempre han ido dirigidas a evaluar aquellos aspectos que contenían los programas que pretendían mejorar la comprensión y casi nunca se han dirigido a evaluar procesos. En efecto, la evaluación del lenguaje y, en concreto, del lenguaje escrito, se ha centrado casi exclusivamente en el registro de respuestas proporcionadas respecto de algún test estandarizado o tarea específica de comprensión lectora. Durante mucho tiempo, muchas de estas pruebas se han dirigido especialmente a valorar el producto de la comprensión (Alonso Tapia y Cols. 1992), a veces con tareas que primaban el recuerdo o estaban exclusivamente centradas en la detección de fallos o errores en las actividades propuestas. Sin embargo, se ha descrito que la simple detección de los errores cometidos por parte de los alumnos en este tipo de tareas, parece prestarse sólo a una descripción de los síntomas concretos que se perciben en el alumnado con dificultades en el área del lenguaje escrito. Se trataría por consiguiente de una evaluación superficial, que, si no se ve complementada por un estudio más exhaustivo del problema, dirigido más a los procesos implicados en la comprensión y no tanto al producto final de la misma, será de poca utilidad en la práctica (Solè, 1992; Alonso Tapia y Carriedo, 1996; Moreno y García, 2000).

Los errores producidos por los alumnos son, sin duda, una poderosa fuente de información y su utilización puede ir más allá del recuento de éstos para extraer conclusiones guiadas por su número final. En efecto, las prácticas encaminadas a favorecer la revisión y detección de los propios errores por parte del alumno, con el apoyo necesario, parecen sacar partido a la comisión de fallos en el proceso de lectura. Otra forma de aprovechar la comisión de errores en tareas de evaluación es la delimitación de problemas, a partir de ahí será necesario poner en marcha una evaluación de procesos implicados en la lectura para poder concretar la problemática y posteriormente realizar la acción educativa adecuada. Es importante hacer hincapié en que, desde esta perspectiva, evaluación e intervención son procesos interdependientes. La evaluación no es tal evaluación si no va encaminada a una intervención eficaz y, a su vez, la intervención debe ser continuamente revisada a la luz de una evaluación continua y, en su caso, modificada; pues bien, difícilmente podemos sostener esto si el proceso de evaluación, en su forma y contenido, se desconecta del proceso de intervención y se centra principal o únicamente en la detección de errores o problemas.

Unido a lo anterior, o más bien ahondando en el tema, la tendencia seguida mayoritariamente por los profesionales de la Psicología durante mucho tiempo ha sido

seguir este modelo de evaluación que se apuntaba, basado en pruebas estandarizadas. Éstas, aunque pueden ser de utilidad para aproximarnos al problema o como complemento para verificar alguna hipótesis, generalmente aportan poca información aprovechable para la intervención directa, máxime cuando la gran mayoría de estas pruebas no han sido construidas para ser analizadas cualitativamente, aún cuando este análisis cualitativo fuera realizado de forma complementaria. Una revisión de algunas de estas pruebas se puede encontrar en Defior y Ortúzar (1993); Jiménez, Rodrigo, Ortiz y Guzmán (1999) y Moreno y García, (2000). Entre estas pruebas estandarizadas cabe destacar algunas recientes que, como la Batería de Evaluación de los Procesos Lectores PROLEC (F. Cuetos, B. Rodríguez y E. Ruano, 1996), tratan de evaluar algunos de los procesos implicados en la lectura; sin embargo, en el caso de la batería PROLEC, la ausencia de una propuesta de evaluación dinámica de los procesos implicados en la comprensión, la consideración de los procesos metacognitivos en un segundo plano, la ausencia de propuestas de contextualización de los instrumentos de evaluación usados e, igualmente, la de propuestas de evaluación del contexto de aula, hacen que esta prueba no responda, a nuestro juicio, a una completa evaluación de procesos en la comprensión de textos.

Otra prueba de evaluación de procesos que, además, centra su atención sobre la valoración de procesos metacognitivos de supervisión y autoevaluación, y sobre la evaluación del conocimiento estratégico, es la batería SURCO de Alonso Tapia, Nuria Carriedo y Mar Mateos (1992). Esta batería constituye uno de los ejemplos más relevantes de evaluación de procesos metacognitivos en la comprensión de textos. A nuestro entender, sin embargo, la realización individual por parte de los sujetos y la valoración exclusiva del producto de la realización de la prueba, ocasionan que, en la práctica, no se evalúe el proceso de resolución de actividades de comprensión, perdiendo, de esta forma, la posibilidad de acceder al conocimiento consciente del sujeto sobre sus propias estrategias y de valorar procesos de control y regulación de la actividad de comprensión durante la propia actividad. La provocación de estas *experiencias metacognitivas conscientes* (Mora y Aguilera, 2000) contribuiría a facilitar la valoración de las capacidades actuales y próximas del sujeto, así como las ayudas necesarias para potenciar las mismas. No se propone, por tanto, en la batería propuesta por Alonso Tapia y Colaboradores, la evaluación dinámica de los procesos. Al ser pruebas estandarizadas, tampoco se propone la contextualización de los instrumentos de evaluación. Por último, tampoco se propone la evaluación del contexto de aula dentro del proceso de evaluación. En nuestra opinión, sólo con actividades que faciliten y guíen la evaluación dinámica del proceso de resolución de la lectura y de los distintos procesos implicados en la comprensión, podremos llegar al conocimiento profundo de las dificultades de aprendizaje de un alumno y de sus posibilidades de mejora. Parece necesaria, por tanto, la aportación de algunos aspectos que consideramos esenciales y que no se encuentran recogidos en las distintas pruebas de evaluación que suelen utilizarse. Estos aspectos son: *la evaluación del contexto de aula, la contextualización de los instrumentos de evaluación, la evaluación dinámica preferente de procesos metacognitivos y la evaluación del proceso de resolución de las tareas planteadas.*

Sin embargo, durante los últimos años, como apuntábamos, se ha producido un cambio significativo en la conceptualización misma de la comprensión. En su libro *Leer, Comprender y Pensar* (1992), Alonso Tapia y sus colaboradores realizaban una revisión de aquellos aspectos que sería necesario evaluar para acceder a la comprensión de por qué un alumno presenta dificultades en la lectura y de cuáles pueden ser esas

dificultades. Incluye, entre esos aspectos, los relacionados con las ideas y los conocimientos que los sujetos poseen acerca de la naturaleza de la lectura y de lo que determina la comprensión. También incluye la necesidad de evaluar los procesos de supervisión y evaluación de la propia comprensión, haciendo alusión a: (a) la detección de los errores o fallos en la comprensión; (b) la valoración crítica de los contenidos; (c) el proceso de construcción del significado —incluyendo la actividad inferencial del sujeto—; (d) las estrategias para distinguir la información relevante y la información, en definitiva, qué sería idóneo retener en función del propósito de la lectura; y, por último, hace referencia a (e) la necesidad de valorar el producto de la comprensión, es decir, la representación que el sujeto construye a raíz de su lectura, tanto la representación textual de base como la representación situacional.

Otros autores han abogado también por una evaluación de procesos en la lectura, que fuera más acorde con las nuevas conceptualizaciones que se han ido construyendo sobre los procesos de lectura y comprensión y que pudiera dar una respuesta más eficaz a las necesidades planteadas en la realidad escolar. Algunas de estas propuestas incluyen también la evaluación de procesos metacognitivos como herramienta esencial para acceder a la *comprensión* de las dificultades de aprendizaje en los alumnos (Mateos, 1991; Alonso Tapia, Carriedo y Mateos, 1992; Alonso Tapia, Carriedo y González, 1992; Peronard, Crespo y Velásquez, 2000; Peronard, Crespo y Guerrero, 2001; Swanson, 2000; Resing, 2000; Kozulin y Garb, 2001; Repetto, Téllez y Beltrán, 2002)

En línea con todo lo anteriormente expuesto, la evaluación del lenguaje —y en nuestro caso concreto, de la lectura— debe ser funcional, orientada a una eficaz intervención y no tanto a un correcto «etiquetado» del problema. Para ello será necesario un minucioso análisis de las habilidades y los procesos que entran en juego en el proceso de comprensión de la lectura. Se trataría de descomponer un proceso en otros múltiples procesos que, por otra parte, son interdependientes y a la vez multicausantes del problema. En este sentido, pueden contribuir técnicas como la de *saturación* del problema (Mora, 1994; Aguilera y Saldaña, 2003; Aguilera y García, 2004). Esta estrategia dinámica de evaluación (Aguilera y García, 2004) facilitaría una detección fina de los posibles errores o dificultades asociados a factores extra-lingüísticos y, lo que es más importante, de los mecanismos a través de los cuales éstos se producen. Por tanto, la evaluación debe estar centrada en los procesos implicados en la lectura y en el proceso mismo de lectura. Como decíamos anteriormente conviene ver productos, pero lo que más nos interesa es encontrar, *en el camino*, dónde se encuentran los obstáculos en los que tropieza el alumno. Estos obstáculos pueden ser relativos a las capacidades personales del alumnado, a procesos relacionados con el contexto —ya sea escolar o socio-familiar—, y también pueden deberse a procesos derivados de la tarea en sí; es decir, de las características del texto que se debe leer. Debemos encontrar dónde está la dificultad, qué habría que hacer para que mejorara, cómo lo conseguiremos y mediante qué sistemas de apoyo.

La necesidad de una evaluación dinámica de los procesos de comprensión y del proceso de lectura y comprensión

Llegados a este punto, parece claro que la evaluación de los procesos implicados en la lectura y del proceso mismo de comprensión de un texto, llegando a establecer aquellas condiciones optimizadoras del desarrollo de estos procesos, debe contemplar la *interacción* entre quien evalúa y el alumno, mediando el primero entre la tarea

propuesta, por una parte, y las habilidades manifestadas por el segundo, por otra, con objeto de ajustar nuestras tareas o preguntas a la respuesta y realizaciones del alumno. Esta forma de actuar se torna básica para acceder al proceso de resolución, por parte del alumno, de una tarea de lectura determinada y para encontrar vías de solución a las posibles dificultades encontradas. La evaluación de la lectura propuesta en este trabajo se basa en esta interacción entre evaluador y alumno con dificultades, observando y mediando entre él y la tarea propuesta con el fin de obtener la mayor información sobre sus dificultades y sobre cómo mejorar el proceso.

Los evaluadores, ya sean tutores, orientadores, profesores de apoyo o logopedas, asumen una posición activa en las situaciones de evaluación; modifican sus preguntas e incluso los propios instrumentos de evaluación para ir adaptándose a las respuestas que va ofreciendo el alumno. Ofrecen instrumentos o pautas de control y regulación de la actividad de lectura para observar y valorar el grado de incorporación, utilización y aprovechamiento del alumno de esos instrumentos en la resolución de la tarea. Intervienen en el proceso de resolución de la tarea mediando su realización, realizando preguntas o comentarios que se constituyan en instrumentos de control y regulación de la actividad, para evaluar el proceso de resolución y el grado y tipo de apoyo que necesita el alumno. Para ello se proporcionan actividades que supongan la posibilidad de observar y analizar el proceso de resolución de una tarea, implicando la puesta en práctica de procesos cognitivos, metacognitivos y de ajuste personal-social.

Para facilitar la evaluación del proceso de resolución de tareas de lectura, proponemos en el dispositivo de evaluación dinámica EDPL la utilización del vídeo. Este recurso permite observar y revisar cuantas veces creamos oportuno el momento de ejecución de la prueba, observar con más detalle al alumno o alumna, sus reacciones, sus decisiones, recoger su habla espontánea, sus preguntas y respuestas, etc. Facilitará igualmente el registro de las conductas observadas y de la respuesta del alumnado en las hojas de registro y evaluación. Por otra parte, el análisis de las situaciones de evaluación proporcionará a los evaluadores una información valiosa acerca del proceso de evaluación en sí mismo, es decir, acerca de la propia actuación encaminada a detectar dónde se encuentran las dificultades y qué podemos hacer para solventarlas. El análisis de las situaciones de enseñanza-aprendizaje en el aula o de situaciones de interacción motivadas por la evaluación o la intervención con alumnado que presenta dificultades de aprendizaje, se revela como uno de los instrumentos más útiles para mejorar la propia práctica educativa y ajustarla progresivamente a los objetivos planteados (Parrilla, 1996; Mora, 1998).

Por otra parte, y como hemos recogido ya en otro lugar (Navarro, 2004), la *construcción del conocimiento escolar es una función de la ayuda prestada contingentemente a las necesidades educativas del alumno* (Luque, Ortega y Cubero, 1997). La ayuda ajustada —en función de la situación de aprendizaje de que se trate, del alumno o del dominio específico— del maestro o compañero más aventajado, media el acceso del alumno menos aventajado en el contenido de que se trate, a unos determinados contenidos culturales. Facilitará, en mayor o menor medida, la tarea de construcción personal a la que se enfrenta. En función de esto, el desarrollo potencial del aprendizaje dependería de dos funciones complejas: (a) por un lado, el desajuste óptimo que debe existir entre la tarea de aprendizaje propuesta y el nivel de funcionamiento cognitivo o socio-moral actual; es decir, es necesario, como se ha dicho anteriormente, situarse en la zona de desarrollo próximo del alumno. Esto exige del maestro conocer el funcionamiento

cognitivo actual del alumno y también otros aspectos más relacionados con sus intereses y motivación, que determinan en gran medida la disposición necesaria de un alumno para aprender un contenido de forma significativa; y (b) por otro, el ajuste de la ayuda del maestro o del alumno más aventajado. Este punto exige el anterior. En este sentido, la posibilidad de que algunos alumnos tutoricen a otros —lo que se ha visto respaldado por la investigación en numerosas ocasiones y especialmente en aquellas referentes al incremento de desarrollo cognitivo (Mora, 1998)— parece idónea dado que, en principio, no debería existir una gran diferencia en el funcionamiento cognitivo entre tutor y tutorizado, lo que sitúa al primero, de una forma espontánea, en la zona de desarrollo próximo del segundo, hecho que no se da entre maestro y alumno, entre los que la diferencia de funcionamiento se hace muy notable.

La labor educativa tiene mucho que ver con proporcionar experiencias de aprendizaje adecuadas, provocando situaciones de interacción en las que el alumnado ponga en juego sus capacidades y recursos actuales para poder mejorarlos. A este respecto cabe destacar que la provocación de la *experiencia metacognitiva consciente*, a la que ya hemos aludido, (Mora y Aguilera, 2000); es decir, aquella experiencia en la que, bien se reflexiona conscientemente sobre los propios procesos mentales, el propio modo de pensar o de enfrentarse a los problemas, o sobre el conocimiento de las propias limitaciones o habilidades; o bien se están poniendo en práctica conscientemente actividades de autorregulación, como la planificación, la supervisión o la evaluación de los propios procesos cognitivos o las estrategias seguidas; constituye quizás la forma idónea de generar actividad mental constructiva por parte del alumnado y, por tanto, desarrollo. Es precisamente durante esta *actividad metacognitiva consciente* cuando, a nuestro juicio, tiene mayor sentido la evaluación del desarrollo potencial del sujeto en cualquier ámbito del aprendizaje, también con respecto a la lectura y la comprensión de textos. Esto se debe, como hemos tenido oportunidad de comprobar, a la relevancia de los procesos metacognitivos en la explicación del rendimiento cognitivo general, de los procesos de comprensión, y también de la adaptación social del individuo (Nisbett y Shucksmith, 1987; Campione, Brown y Ferrara, 1987; Mora, 1991, 1998; Martín, 1999; Kozulin, 2000; Carlson y Wiedl, 2000).

En este sentido, aunque en recientes trabajos relacionados con la evaluación se han comenzado a tener en cuenta los procesos metacognitivos, a los que ya aludía Vygotski en sus escritos sobre la atención educativa a los niños con déficit intelectual (Vygotski, 1931/1997), no parece haberse otorgado a estos la importancia que, según el propio Vygotski en un principio, y, posteriormente, numerosas investigaciones estrechamente relacionadas con el concepto de la evaluación del desarrollo potencial del aprendizaje, han podido poner de manifiesto (Flavell, 1976; Brown, 1980; Nisbett y Shucksmith, 1987; Campione, 1989; Campione y Brown, 1987; Mora, 1998). La capacidad de reflexión consciente sobre nuestros propios pensamientos y conocimientos, nuestras experiencias, habilidades o actitudes, así como la capacidad para controlar y regular estos mismos procesos, se constituyen en un aspecto central si hablamos de adaptación flexible al medio existente y a los cambios que se suceden o se introducen en éste y, por tanto, si queremos abordar la posibilidad de originar o facilitar modos más óptimos —ya sea a niveles estructurales o funcionales— en la forma de enfrentarse a estos cambios. Para Vygotski era precisamente sobre estos procesos de control y regulación cognitiva, sobre los que había que intervenir para modificar la capacidad de un sujeto. Esto cobraba aún más relevancia cuando se trataba de un sujeto con deficiencia mental. Consecuentemente, parece obvio que una evaluación dinámica de estos procesos

metacognitivos (Saldaña, 2001) puede otorgar información valiosa por estar centrada en el proceso mismo de aprendizaje y por orientarse a evaluar procesos superiores que pueden controlar y regular las propias habilidades de pensamiento y de interacción social.

Si, como hemos podido comprobar, el aprendizaje se da en contextos de interacción, es necesario tener esto en cuenta, tanto a la hora de intervenir pedagógicamente como también a la hora de evaluar. Este nuevo enfoque de las situaciones de aprendizaje llevó de la mano, como hemos visto, un nuevo modo de afrontar las situaciones de evaluación (Vygotski, 1934/1995; Minick, 1987; Lidz, 1995). No basta con establecer el nivel evolutivo en relación a lo que un niño es capaz de realizar por sí mismo y de forma autónoma, sino que es necesario evaluar qué es capaz de realizar con la ayuda de alguien más experto (Vygotski, 1934/1995); esto permitirá establecer con mayor validez qué funciones se encuentran en periodo de maduración y cuál sería el mejor modo de ofrecer ayuda para llegar a su consecución por parte del alumno. Esta «medida» de las capacidades intelectuales de los niños suponía para Vygotski una indicación de la dinámica de los procesos cognitivos, sin duda, más útil que la que podían ofrecer pruebas psicométricas con la indicación de «edad mental», o que la que, hoy día, pueden ofrecer muchas de las pruebas estandarizadas comercializadas.

En definitiva, la aparición de la evaluación dinámica, como hemos tenido la oportunidad de ver, ha estado relacionada desde el comienzo, con la revisión de supuestos teóricos sobre el desarrollo de la inteligencia y el modo idóneo en que ésta podía medirse. En este sentido destacamos algunos de estos elementos que suponen una revisión de esos postulados teóricos basados en la medición cuantitativa de capacidades manifiestas, que reflejaría a su vez las aptitudes internas; en la ejecución individual y autónoma como la mejor forma de evaluar dichas capacidades; y en la clasificación de los sujetos en función de los resultados en pruebas estándar. Estos postulados mencionados sobre la evaluación, como hemos descrito anteriormente en este trabajo, también han presidido la evaluación de la lectura a lo largo de muchas décadas. En definitiva, ponemos énfasis en una evaluación dinámica de procesos en la lectura porque: (a) se trata de evaluar un proceso, en este caso el de la lectura, no un resultado; por tanto no se trata de un elemento estático, sino dinámico por naturaleza; (b) se trata de evaluar los procesos de lectura para mejorarlos. Se busca descubrir aquellos elementos que pueden estar interfiriendo en una adecuada comprensión del texto escrito, para optimizar los esfuerzos de apoyo; (c) la lectura es un proceso susceptible de modificación por el aprendizaje y esto debe ser tenido en cuenta a la hora de la evaluación; y (d) se trata igualmente de descubrir aquellas condiciones (material y apoyo necesario) bajo las cuales los sujetos obtienen mejores resultados en su lectura.

Una perspectiva metacognitiva. Nuestra propuesta de evaluación dinámica de procesos metacognitivos en la comprensión de textos

¿Cómo valorar finalmente los aspectos que se recogían en la Figura 1 y su incidencia en el acceso a la comprensión del texto? Ha sido para nosotros una constante preocupación el discernir ante cada actividad que se analizaba o elaboraba, qué habilidades metacognitivas se ponían en juego al tener que realizarla o cómo incidían éstas en un adecuado proceso de resolución de la misma. Esto nos ha llevado a un ejercicio de reflexión y discusión sobre cada una de las actividades, del cual han surgido las que se citan en este trabajo. No están exentas de limitaciones; es más, nuestra

intención ha sido desde el principio ir mejorándolas a medida que encontrábamos elementos que completaban las ya existentes o introducían aspectos no considerados hasta entonces. De esta forma se han ido modificando algunas de las actividades y tareas que se eligieron o diseñaron al comienzo.

Hemos elegido una doble forma de evaluar estas habilidades metacognitivas implicadas en el proceso lector; por una parte, utilizando algunas actividades específicas que tratan de poner de manifiesto la calidad de la utilización de este tipo de habilidades. La primera de estas actividades específicas consiste en la detección de incoherencias en un texto, tratando de elicitar la puesta en práctica de procesos de supervisión de la propia comprensión. Las incoherencias introducidas en los textos son de tipo léxico, sintáctico y semántico. Así mismo, con la mediación necesaria, el alumno deberá recomponer el sentido de los textos, para dar solución a las demandas efectuadas sobre los mismos por los evaluadores en relación a aspectos estructurales o de contenido; en otra de estas actividades los/as alumnos/as deben comparar dos textos en función de la dificultad de su comprensión, provocando un análisis sobre dónde se encontraban estas dificultades, sobre los elementos estructurales que facilitan o dificultan la comprensión y sobre cómo solventar las dificultades encontradas; y, por último, en otra tarea se pide a los/as alumnos/as que expliciten, mediante la cumplimentación de un cuestionario de opción múltiple, sus metaconocimientos acerca del significado y sentido de la lectura, así como sobre las estrategias que pueden usarse para comprender los textos.

Por otra parte, se recoge información sobre la puesta en práctica de estas habilidades metacognitivas durante la realización de las diversas tareas propuestas, mediante la observación sistemática y la entrevista metacognitiva. Se trata de aprovechar los momentos de resolución de las tareas por parte del alumnado para observar y realizar preguntas y sugerencias que nos orienten acerca del conocimiento que tienen los sujetos sobre su proceso de comprensión y sobre las propias estrategias, y que nos den información sobre la puesta en práctica de procesos de autorregulación. En el desarrollo de cada una de las tareas, en el apartado correspondiente a *Mediación y evaluación de los procesos metacognitivos*, se recogen algunas de estas preguntas y sugerencias a modo de orientación. Hemos tratado de compensar, utilizando más de una forma de evaluación de las habilidades metacognitivas implicadas en el proceso lector, las carencias que, según la investigación sobre la evaluación del proceso de comprensión —como hemos tenido oportunidad de comprobar en el capítulo 1—, ofrecía la entrevista metacognitiva como única fuente de evaluación, debido a las posibles diferencias entre los procesos en sí y la verbalización que de ellos podía hacer el alumno (Garner, 1987). Así pues, optamos por estos tres elementos: *observación* y *entrevista metacognitiva*, por un lado, y *actividades específicas*.

La evaluación del contexto de aula

En otro lugar (Navarro, 2004) hacíamos mención a la relevancia del contexto en la construcción del conocimiento escolar. Analizábamos una de las premisas relativas al constructivismo en los procesos educativos, que fue a su vez recogida en un clarificador trabajo por Luque, Ortega y Cubero (1997); la premisa en cuestión hacía referencia a que *el contexto influye en la construcción de los conocimientos y capacidades porque da sentido a la experiencia*. Comentábamos que no hacíamos alusión solo a un contexto material, social o geográfico, sino que hacíamos también referencia al contexto discursivo que se hace presente en las interacciones y condiciona la comprensión y la

actividad de los sujetos que participan. Pues bien, si el contexto influye en la construcción de los conocimientos, sin duda debe influir también cuando estos conocimientos (o igualmente si hacemos referencia a procedimientos o actitudes) deben ser mostrados y puestos en práctica por parte de los alumnos y las alumnas de cara a una evaluación. Por consiguiente, parece obvio pensar que debemos incidir aún más en la importancia de realizar evaluaciones que tengan en cuenta dónde y de qué forma se han desarrollado los aprendizajes o las capacidades que vamos a valorar (Romero y Lavigne, 2006). Esto es importante, tanto si hablamos de una evaluación que realiza el profesor en su aula para obtener información del grado de aprendizaje que se ha producido durante un determinado trabajo, como si lo hacemos de una evaluación psicopedagógica que realiza el orientador junto con el profesor del aula para determinar posibles necesidades educativas de alumnos o del propio contexto de enseñanza-aprendizaje.

En este sentido parecen ser: (a) la observación sistemática y el análisis de situaciones naturales en las que los alumnos, dentro de su aula, deben desplegar sus conocimientos, habilidades o actitudes; (b) el análisis de las interacciones (profesor-alumno, alumno-alumno), con el modo, objetivos e intencionalidad con la que éstas se llevan a cabo; (c) el análisis de las tareas que de forma efectiva se llevan a cabo (contenidos, estructura física del aula en relación a esa tarea, cómo se evalúa la misma, ayudas o apoyo que se prevé para el acceso de todos/as los alumnos a los contenidos propuestos...); (d) el análisis de la estructura organizativa del aula, es decir, modo de organizar un trabajo determinado durante un tiempo determinado (distintos tipos de comunicación, secuencias participativas, normas y valores establecidos, clima del aula...); y (e) el análisis de las ideas, pensamientos, conocimientos y creencias del profesor/a en relación a los procesos educativos que acontecen en el aula; aquellos aspectos significativos y determinantes para obtener información relevante acerca de los procesos estudiados y poder realizar propuestas de mejora sensibles al contexto educativo donde esos procesos tienen lugar. Obviamente, la evaluación del desarrollo potencial del aprendizaje debe realizarse teniendo en cuenta los aspectos destacados anteriormente; no parece tener mucho sentido una evaluación de los procesos cognitivos en maduración y de las condiciones contextuales que favorecerían la mejora de esos procesos, de modo ajeno al contexto (material, social, discursivo...) donde habitualmente se trabajan y se ponen en práctica dichos procesos.

Cuando hablábamos de la conveniencia de realizar una evaluación dinámica, destacábamos una serie de puntos que creemos fundamentales para desarrollar una evaluación de procesos implicados en la lectura. Entre estos puntos se encuentran aquellos que hacen referencia a la susceptibilidad de modificación de la comprensión lectora por medio del aprendizaje, y al descubrimiento, a través de la evaluación, de aquellas condiciones (material y apoyo necesario) que mejorarían el proceso de lectura en un alumno. De estas premisas se desprende también la relevancia que cobra en nuestra propuesta la *evaluación del contexto*. En este sentido, hacemos referencia principalmente al contexto de aula, al ser habitualmente el espacio natural donde se desarrollan las actividades encaminadas a la mejora intencional de las habilidades lingüísticas relacionadas con la lectura. Sería idóneo conocer y evaluar el contexto familiar; sin embargo esto no siempre es posible, como es sabido. De cualquier forma, toda información relevante recogida del contexto familiar puede ser de vital importancia a la hora de plantear nuestras propuestas de mejora.

La evaluación del contexto no siempre es tarea fácil. A veces se puede tener la sensación de que el realizar determinadas tareas encaminadas a evaluar el contexto de aula, donde, por otra parte, se presentan las dificultades en el aspecto que estemos tratando (en este caso comprensión lectora), no se relaciona con la posible mejora de los problemas que «plantea un determinado alumno». Sin embargo, puede ser fácil hallar acuerdo en que, tras observar minuciosamente aspectos como: (a) las interacciones que tienen lugar entre los «actores» del proceso (profesor-alumnos o alumnos-alumnos); (b) los agrupamientos, que, bien de forma espontánea o guiada, conforman una parte fundamental del escenario donde tienen lugar las actividades de enseñanza y aprendizaje y que, a su vez, condicionan en mayor o menor medida, las interacciones que tienen lugar; (c) las actuaciones metodológicas encaminadas a la consecución de los objetivos educativos; (d) las actividades concretas que debe realizar el alumnado; etc., —todos ellos procesos que tienen lugar en un aula y que, por lo tanto, no pueden evaluarse fuera de ella—, no siempre éstos favorecen la adquisición o el avance de determinados alumnos en la comprensión de textos.

Parece necesaria la adopción de un enfoque ecológico-sistémico, basado en los trabajos de Bronfenbrenner (1987), donde es relevante la valoración de cada uno de los elementos presentes en el aula —y a ser posible fuera de ella— que estén relacionados con la situación de dificultad que se manifiesta en uno, dos o quizás más alumnos; y donde también hay que tener en cuenta que cada uno de estos elementos influye y se ve influido por los demás. Igualmente relevante nos parece la adopción de perspectivas de análisis más *explicativas* del desarrollo educativo escolar en contextos significativos de aprendizaje, como son las aproximaciones histórico-culturales de Michael Cole y su grupo de investigación (Cole y D`Andrade, 1982; Moll (1989). Así, nuestra actuación docente a la hora de promover las actividades de aprendizaje de la lectura —por ejemplo—, tendrá una indudable influencia en el alumnado; desde el método concreto que elegimos: fonético, global o mixto, a la opción por desarrollar el proceso de enseñanza y aprendizaje de la lectura sin el apoyo de un método concreto; desde la relevancia que otorgamos a la dotación de significatividad desde el comienzo de todo aquello que leemos, a la consideración, por el contrario, de la decodificación como el proceso central en el aprendizaje de la lectura; desde opciones que contemplan la lectura en voz alta de cuentos u otras lecturas, interactuando con los alumnos y haciéndolos partícipes de la actividad, a concepciones más basadas en la realización sistemática de ejercicios de reconocimiento de grafías; etc. Todo ello tendrá, sin duda, una repercusión en la adquisición y asimilación que de este contenido hagan todos los alumnos y las alumnas del aula, así como en las ideas que sobre la lectura, su sentido y significado, elaboren durante el proceso.

De este modo, la evaluación del contexto de aula debería permitirnos recoger información sobre los procesos educativos que tienen lugar dentro de ella, y poder elaborar propuestas metodológicas a nivel de aula que mejoren estos procesos con relación al tema que nos ocupe. Puede tratarse de optar por la diversificación del método de aprendizaje de la lectura y la escritura en función de las características de nuestro alumnado; o de realizar tareas que refuercen lo ya aprendido por otros alumnos, con aquellos que han presentado más dificultades; o de promover otros agrupamientos que, beneficiando especialmente a los alumnos con más dificultades, benefician de igual modo a aquellos que hacen de modelos o deben explicar cómo se realizan determinadas actividades y apoyar a sus compañeros; etc. Algunas de estas propuestas metodológicas concretas para poner en práctica, tanto a nivel de aula como referidas al alumnado con

especiales dificultades en la lectura, quedan recogidas en el volumen de Anexos, en el que se analiza la propuesta de intervención del dispositivo dinámico EDPL.

Esta evaluación del contexto se concreta en nuestra propuesta mediante dos instrumentos: *la observación dentro del aula* —tanto por parte del profesor de la misma, como, si es el caso, por parte del orientador, profesor de apoyo, logopeda, o cualquier otra persona que esté participando en la evaluación— y *la entrevista con el profesor*, en caso de que participe de la evaluación una persona externa al aula. En la observación dentro del aula, para facilitar tanto nuevas observaciones del mismo proceso, como el posterior registro, proponemos, al igual que hicimos anteriormente cuando hablábamos de la evaluación dinámica, la utilización del video. Esto nos permitirá probablemente extraer más información relevante sobre lo acontecido en el aula con relación al trabajo sobre lectura, al mismo tiempo que revisar nuestro propio proceso de evaluación cuando lo creamos oportuno o permitimos la valoración de las sesiones por otros profesionales para, de esta forma, obtener una mayor fiabilidad de nuestras observaciones.

Evidentemente, nos interesa sobre todo observar actividades relacionadas con los procesos de lectura. Por esto hemos adaptado de Sánchez, Rosales y Suárez (1999) un procedimiento de análisis de la interacción profesor-alumnos en la comprensión de textos. Con la aplicación de este instrumento, pretendemos aproximarnos a lo que efectivamente se realiza en el aula con relación a la comprensión lectora, para poder ofrecer esta información al profesorado y, juntos, si es oportuno, elaborar propuestas de mejora para todos o para algunos alumnos.

Por otra parte, la entrevista nos permitirá principalmente conocer el trabajo que se viene haciendo con aquel o aquellos alumnos con dificultades en la lectura, el punto de vista del profesor sobre las causas que motivan tales dificultades, así como otros aspectos que pudieran ser relevantes con relación al alumno o el aula. Esta entrevista puede realizarse en cualquier momento del proceso de evaluación. Quizás lo más conveniente sea iniciar el proceso obteniendo información de aquella(s) persona(s) que más conoce(n) y que más tiempo trabaja(n) con el alumno o alumnos evaluados. La información procedente de él/ella será de gran ayuda, pues probablemente, a través de sus descripciones, el problema adquiera límites más concretos, contribuyendo a ubicarlo mejor en algún lugar de nuestra *galaxia* de procesos implicados en la lectura. Si la información recogida en esta primera entrevista es demasiado vaga o imprecisa, o se presta a confusiones, no debe suponer ninguna traba, pues, coherentemente con nuestra propuesta, podemos comenzar por el análisis de cualquiera de los procesos implicados en función de la primera hipótesis que hayamos construido con la información de que dispongamos. Sería la evaluación del proceso de resolución de las tareas y de los resultados que se vayan obteniendo, los que marquen el hilo conductor; así, será a posteriori de la realización de una tarea determinada cuando se sabrá cuál es la que vendrá a continuación, en función de los datos obtenidos.

Para guiar esta entrevista ofrecemos algunas cuestiones que pueden servir de orientación en función de la información que nos interesa recoger. Algunas de estas cuestiones han sido recogidas y desarrolladas en el Cuestionario EIDAL, sobre ideas y expectativas del profesorado con relación a las dificultades de aprendizaje de la lectura; este cuestionario ha formado parte de los instrumentos aplicados de forma complementaria en nuestro estudio.

Cuadro 22. Guía para la obtención de información relativa al contexto de aula

- *qué dificultades manifiesta el alumno (o grupo de alumnos) en su lectura
- *qué origina y mantiene, según el profesor, estas dificultades, qué elementos pueden estar influyendo en el problema: el propio alumno (su motivación, actitud, capacidad desarrollada hasta ahora...), los textos que debe leer (excesiva dificultad, inadecuación del contenido...), el contexto (escolar o familiar)
- *en qué momentos manifiesta estas dificultades: siempre, con cualquier texto, en cualquier situación de aprendizaje, o están ligadas a un área concreta o a situaciones determinadas (por ejemplo cuando debe exponer sus trabajos o leer en voz alta, etc.)
- *qué actuaciones se han llevado a cabo para paliar estas dificultades
- *cuál ha sido el resultado de estas intervenciones
- *si existen programas específicos que se estén trabajando para favorecer la comprensión lectora en general

Contextualizar los instrumentos de evaluación

La evaluación del proceso de lectura mediante tareas que no difieren mucho de las que el alumnado realiza habitualmente en su aula, realizada en el contexto donde habitualmente se desarrollan las actividades encaminadas a la mejora de estos procesos (Romero y Lavigne, 2006), contribuye notablemente a incrementar la validez ecológica de los resultados de este proceso de evaluación, lo que a su vez se relaciona con la mayor posibilidad de generalizar las conclusiones obtenidas en la evaluación; así mismo, contribuirá de manera importante a conectar los procesos de evaluación e intervención posterior para la optimización de la habilidad lectora (Navarro, 2004).

Por tanto, nos ha parecido importante contextualizar, en la medida de lo posible, algunas tareas que le son propuestas al alumnado. De esta forma, algunas actividades no están cerradas en esta evaluación, sino que debe ser la persona que ponga en práctica estos instrumentos, la que termine de elaborar las mismas, sirviéndose, en algunos casos, de textos cercanos y significativos para los alumnos, o del conocimiento de palabras familiares para éstos. Bajo nuestro punto de vista, es éste un elemento crucial en nuestra propuesta; se trata de realizar una evaluación de la comprensión lectora que se actualiza cada vez de forma distinta ante cada alumno. Esto, sin duda, complejiza un tanto la labor de quienes deben valorar este aspecto: profesores y/o orientadores; pero nos da cierta garantía sobre la validez de los resultados.

Ante el hecho de que, a primera vista, pudiera parecer excesivo el número de actividades que finalmente hemos considerado relevantes para la evaluación de la comprensión lectora, querríamos hacer dos comentarios: por un lado hemos de decir, como ya quedó expuesto en otro trabajo (Varios Autores, 1998), y también en este mismo, que la lectura es un proceso enormemente complejo, que requiere poner en

juego una serie de elementos y acciones cognitivas, metacognitivas, emocionales, conductuales y de contexto, que, de forma inter-relacionada, van a producir el hecho subjetivo de la comprensión de un texto escrito; bajo nuestro punto de vista, este proceso complejo no puede ser evaluado de forma simple si pretendemos recoger la información suficiente para poder detectar de una forma lo más precisa posible aquellos elementos que dificultan la comprensión y poder así contribuir a la elaboración de propuestas de mejora.

Por otra parte, tal y como recogíamos en la figura 1, nuestra propuesta parte de una *galaxia* de elementos y acciones implicadas en la comprensión lectora. La peculiaridad de este modo de representar estos elementos implicados en el proceso estudiado, consiste en que otorga la posibilidad al evaluador de «entrar» a valorar cualquier proceso que crea oportuno, sin necesidad de hacerlo de una forma lineal (recuérdese que la lectura es un proceso interactivo, en el que elementos que formarían parte de procesos de ajuste personal-social, o implicados en la integración textual, o de integración del texto con los conocimientos y experiencias previas, pueden estar influyendo en la comprensión, incluso antes de que se pongan en marcha mecanismos más ligados a procesos perceptivos o de asociación G-F). Así mismo, no sería necesario de este modo, realizar todas y cada una de las actividades propuestas, pudiéndose optar por realizar sólo algunas, pertenecientes a un determinado procesamiento descrito, o, sencillamente valorar tan sólo un procesamiento si se posee información anterior relativa a la no existencia de dificultades en los demás aspectos. Relacionado con lo anterior, otro aspecto que debe tenerse en cuenta para matizar un poco más el hecho del número de actividades propuestas, es el de que determinados procesos que quedan recogidos en el esquema general (capítulo 6), no necesitarían de una actividad específica para su evaluación, aprovechándose para ese fin alguna de las tareas propuestas en las que también se valoren estos procesos.

La propuesta de evaluación dinámica contenida en el dispositivo EDPL pretende ser lo más *completa* y *útil* posible. Se pretende valorar, no sólo las capacidades actuales y manifiestas en cuanto a la comprensión de textos, sino, sobre todo, qué elementos del proceso dificultan o interfieren la comprensión y cómo podemos contribuir a mejorar la misma. Hemos tratado de contemplar algunas actividades o tareas cuyo objetivo sería «afinar» lo más posible en nuestra evaluación, una vez que hemos detectado en qué elementos del proceso se encuentran principalmente los problemas. De este modo, se proponen actividades para realizar en el caso de que se produzcan dificultades significativas en otras pruebas. En el *esquema de las actividades de evaluación dinámica*, recogido en el capítulo 6, donde se describe y analiza el dispositivo EDPL, estas actividades mencionadas anteriormente, se encuentran bajo —o a la derecha de— una línea discontinua. Igualmente, en la *galaxia de actividades de evaluación*, que se encuentra igualmente en el citado capítulo, estas actividades aparecen en recuadros enmarcados con líneas discontinuas. Por otra parte, en la metodología de aplicación de cada una de las actividades de evaluación propuestas en el dispositivo, queda recogida su utilización de una forma más clara y precisa.

3.2 La evaluación dinámica de procesos implicados en la lectura. Objetivos de la presente investigación

Aunque los objetivos de nuestro estudio pueden haber quedado ya expuestos al esbozar las bases teóricas y metodológicas sobre las que se ha sustentado el trabajo desarrollado, recogemos a continuación de una forma más concreta y operativa estos mismos objetivos que nos han guiado en la práctica a la hora de la elaboración y el desarrollo del estudio. Han sido los siguientes:

1. Comprobar y analizar la eficacia e idoneidad de aplicar instrumentos y procesos de evaluación dinámica en la lectura, de centrarse en una evaluación de procesos y de analizar de forma preferente los procesos metacognitivos implicados en la comprensión, para la determinación de propuestas que mejoren la capacidad de lectura y comprensión en alumnos y alumnas con dificultades de aprendizaje.
2. Extraer y elaborar un conjunto de *patrones metodológicos de evaluación dinámica utilizados en la mediación*, con el objetivo de orientar y optimizar el proceso de enseñanza y aprendizaje de la lectura en sujetos con especiales dificultades.
3. Comprobar, así mismo, que la aplicación del dispositivo de Evaluación Dinámica (EDPL) y la intervención propuesta, centrada en los procesos metacognitivos implicados en la comprensión (IPML), puede llegar a mejorar el rendimiento en lectura de este alumnado, así como su nivel de ajuste personal-social o su capacidad de resolución de una prueba de inteligencia general.
4. Comprobar y analizar la valoración otorgada por el profesorado colaborador con relación a la información obtenida en el proceso de aplicación del dispositivo EDPL, relativa al establecimiento de propuestas de intervención educativa que optimicen el desarrollo del proceso de comprensión.
5. Comprobar y analizar elementos relativos a la fiabilidad y validez de predicción del dispositivo EDPL con relación a la valoración del profesorado sobre el rendimiento académico.
6. Analizar en qué medida los resultados de la aplicación del dispositivo de evaluación dinámica EDPL, ofrecen una información adicional con respecto a la ofrecida por las pruebas estáticas de comprensión de textos, ajuste personal-social o inteligencia.

Esta información adicional debería concretarse:

- en la posibilidad misma de evaluación del *proceso* de resolución de las actividades de lectura;
- en la obtención de mayor información, y de más calidad, sobre la puesta en práctica de los procesos implicados en la lectura, y especialmente, de los procesos metacognitivos; aspectos también valorados en las pruebas estáticas de criterio;
- en la posibilidad de obtener información acerca de aquellos *patrones metodológicos de mediación* que se relacionaran con mejoras en la comprensión de textos, los procesos de ajuste o el rendimiento cognitivo de los sujetos con especiales dificultades de lectura.

MÉTODO

4. Diseño de investigación e hipótesis planteadas

4.1 El diseño de investigación

El diseño de investigación utilizado ha sido el de *pretest-posttest con grupo control no equivalente*. Se ha comparado por tanto la actuación de dos grupos de sujetos a través de diversas condiciones experimentales con el objetivo de observar y analizar si la introducción del conjunto de instrumentos que conforman el tratamiento en el GE es causa de efectos diferenciales en su actuación post-tratamiento con relación al GC. Se trata de un diseño en esencia cuasiexperimental, en el que los grupos formados no pueden ser del todo equivalentes, ni el control ejercido sobre las condiciones experimentales puede llegar a ser absoluto (Sarriá, 2001). La posibilidad de ejercer el control de las variables potencialmente contaminantes mediante la aleatorización completa de todos los efectivos de la muestra, no es posible en nuestro caso, ya que, como hemos visto, los sujetos han sido seleccionados en función de su pertenencia al grupo de alumnos y alumnas con dificultades especiales en su lectura, y su participación en el estudio ha estado ligada a su vez, a su pertenencia a pequeños grupos que recibían atención educativa especial en las aulas de apoyo a la integración o mediante otros sistemas de refuerzo educativo por parte del colectivo de profesionales que ha colaborado en el estudio. La asignación de los profesores-colaboradores y, con estos, de sus grupos de alumnos participantes, sí ha sido realizada en parte de forma aleatoria, aunque también en función de las posibilidades reales de participación en el Grupo de Trabajo que tendría lugar con objeto de la aplicación y seguimiento del dispositivo de evaluación dinámica. Por otra parte, ante la imposibilidad de un control completo y total de todas las variables intervinientes en el proceso de investigación, se ha tratado de conocer en grado máximo estas variables y de ejercer el mayor grado de control posible con el objetivo de que su influencia en el proceso y en los resultados del estudio se haya visto minimizada (Lukas y Santiago, 2004).

Estas condiciones del diseño elegido, como ha sido expuesto en algunas ocasiones (Mora, 1998; Sarriá, 2001; Lukas y Santiago, 2004), son casi obligadas en la investigación educativa si se quiere optar por diseños de corte experimental y ejercer un cierto grado de manipulación y control de las variables con el objetivo de obtener mayores cotas de validez interna y conclusiones más sólidas con las que poder establecer relaciones causa-efecto entre las variables. Las decisiones tomadas a este respecto, por tanto, se han convertido en las mejores posibles. Como destacaban ya Campbell y Stanley (1966) o Cook y Campbell (1976, 1979), citados en Sarriá (2001), el cada vez mayor interés por el estudio de fenómenos sociales, culturales y educativos en su contexto natural y las limitaciones que, por otra parte, presentaban estos estudios en cuanto a la posibilidad de control completo de las variables que intervenían en estas situaciones, impulsó progresivamente la sistematización de procedimientos cuasiexperimentales de investigación, en los que se intervenía específicamente sobre las situaciones de estudio, ejerciendo una mayor o menor manipulación de la variable independiente o de tratamiento, con el objetivo de observar su efecto sobre otra(s) variable(s), y ejerciendo el mayor grado de control posible sobre los factores que intervienen en la situación. El elemento diferenciador esencial entre un diseño de este tipo y uno experimental en sentido estricto es, como ya hemos apuntado, la

imposibilidad de asignar aleatoriamente los sujetos a los grupos de tratamiento, con lo que deben afrontarse los estudios con la ausencia de equivalencia inicial total entre los grupos de investigación. El desarrollo y la proliferación de este tipo de diseños no han hecho sino incrementar el estudio de sus limitaciones con el objetivo de mejorar los procedimientos de análisis y la validez de las conclusiones obtenidas. En este sentido, se destaca la introducción de técnicas como el grupo control, la aplicación y análisis de pretest, la realización de análisis de covarianzas, teniendo en cuenta las diferencias previas entre los grupos en el análisis y establecimiento de las diferencias post-tratamiento, así como otros procedimientos estadísticos de aplicación durante el análisis de datos, pero que sirven igualmente al establecimiento de condiciones de equivalencia entre los grupos, como la regresión múltiple. No obstante, se hace necesario ser cauto en las inferencias relativas a la causalidad de los efectos comprobados, así como ejercer el máximo control posible sobre los posibles efectos encontrados, que pudieran ser causados por diferencias existentes ya entre los grupos de tratamiento ajenos a la introducción de la variable independiente.

Como hemos ya señalado en otro lugar (Navarro y Mora, 2002), en investigación educativa, el hecho de realizar una selección diferencial de los sujetos en función de que cumplan una determinada característica —en este caso manifestar especiales dificultades de lectura— es, sin duda, deseable a veces; es decir, muchas de las hipótesis experimentales, o cuasiexperimentales, que se formulan en las investigaciones de carácter evolutivo, educativo o social, cuentan deliberadamente con posibles efectos en los resultados del estudio, de una característica previamente seleccionada en los sujetos de la muestra. Es necesario seleccionar previamente determinadas características de un grupo de alumnos si lo que se quiere comprobar y analizar es el efecto de la introducción de un determinado tratamiento precisamente en ese grupo de alumnos con esas características. Esto suele ser una característica diferencial de la investigación educativa. En nuestro caso, la selección de la muestra —en la que nos detendremos más tarde— ha venido dada principalmente por el nivel educativo que cursaban los alumnos y por la presencia de dificultades de aprendizaje en la lectura. En todo caso, hemos de concluir que los controles ejercidos sobre las variables y el diseño utilizado, responden a la mejor opción posible y tratan de minimizar los efectos negativos sobre la validez interna de la prueba.

Se ha descrito, por otra parte, que la composición de muestras a raíz de grupos clase, en los que, como hemos visto, pueden existir diferencias en cuanto al nivel de rendimiento previo en los aspectos que nos disponemos a evaluar, diferencias relacionadas con la presencia de profesores distintos, que, a su vez, pueden tener estilos educativos que, al menos en parte, difieran; diferencias en cuanto a nivel socio-cultural de procedencia; etc., puede constituir una ventaja, más que un inconveniente (Mora, 1998). La consideración en nuestra investigación, de grupos-clase distintos para la evaluación de la aplicación piloto, y la intención de recoger el máximo de información posible sobre las circunstancias pedagógicas, socio-culturales o de cualquier otra índole, ya sea sobre el grupo completo o que afecten a algún alumno en especial, nos ha permitido establecer descripciones y análisis con relación a los factores estudiados, al tiempo que realizar comparaciones entre grupos, que han respondido mejor a la realidad escolar y a la composición real de los grupos. Han podido abordarse de esta forma la validación de nuestras ideas previas e hipótesis en función de las características de las submuestras que han compuesto la muestra total o incluso de un grupo de sujetos que cumple una determinada característica. Podríamos decir que los grupos de tratamiento, a los cuales se han aplicado las pruebas elaboradas o el dispositivo de evaluación dinámica para

observar los resultados, son tantos como grupos queramos considerar en función de que cumplan una característica que sea relevante para el estudio (tener o no dificultades de aprendizaje, pertenecer a un Centro de Actuación Educativa Preferente, participar de programas de Diversificación Curricular, encontrarse escolarizado en cuarto curso de educación primaria o en sexto, etc.). Por tanto, en lo que respecta a los distintos subgrupos contemplados en la muestra, el análisis que se efectuará en función de las características y circunstancias de aplicación de la actividad de evaluación, y en función a su vez, de las observaciones y comparaciones que se realizarán con relación a esas circunstancias de aplicación para un grupo de sujetos, un grupo clase o un alumno de forma individual, debe considerar distintas observaciones, tanto en la fase pretest como en la fase post-tratamiento. Esto puede quedar esquematizado de la siguiente forma, con relación a cada una de las submuestras y características principales contempladas:

Fase pretest		Fase postest
O1 O2 O3 →	Con relación al nivel educativo	← O4 O5 O6
O7 O8 O9 →	Con relación a la discapacidad	← O10 O11 O12
O13 O14 O15 →	Con relación a las dificultades de apr.	← O16 O17 O18
O19 O20 O21 →	Con relación a la deprivación socio-cul.	← O22 O23 O24
O25 O26 O27 →	Con relación a si recibe apoyo	← O28 O29 O30

Por otra parte, como puede observarse en el esquema, en nuestro estudio se ha contemplado la utilización de varias pruebas de evaluación con el objetivo de proporcionar una información más amplia y compleja sobre los procesos evaluados. De esta forma, se produce una cierta triangulación de la información que es obtenida a través de las distintas pruebas y que es completada con la información de carácter complementario y cualitativa proporcionada por el profesorado participante y colaborador a través de cuestionarios y plantillas de recogida de datos. La diversificación de las observaciones efectuadas proporciona de entrada un control experimental extra en función de la posibilidad de cotejar los datos ofrecidos por las distintas pruebas. En el caso de que estos datos confluyan en una determinada dirección, apuntada previamente en nuestras hipótesis de trabajo, las conclusiones tendrán un mayor grado de validez; por el contrario, resultados altamente divergentes deben hacer examinar con una mayor cautela los resultados finales y aconsejan prudencia a la hora de establecer conclusiones. Al mismo tiempo, resultados divergentes en las distintas observaciones de un fenómeno, pueden enriquecer y matizar los resultados por cuanto pueden llegar a informar sobre grados diversos de respuestas otorgados en función de condiciones o métodos de valoración diversos. El hecho de contemplar distintos subgrupos con características específicas que se consideran relevantes en el análisis del proceso y de los resultados, y el uso de distintas pruebas de evaluación en las fases inicial y final del estudio, aconsejan un diseño como el siguiente:

O1 O2 O3	X1	O4 O5 O6
O7 O8 O9	X2	O10 O11 O12
O13 O14 O15	X3	O16 O17 O18
O19 O20 O21	X4	O22 O23 O24
O25 O26 O27	X5	O28 O29 O30

O1 O2 O3	X1	O4 O5 O6
O7 O8 O9	X2	O10 O11 O12
O13 O14 O15	X3	O16 O17 O18
O19 O20 O21	X4	O22 O23 O24
O25 O26 O27	X5	O28 O29 O30

En definitiva, este diseño nos ha permitido efectuar comparaciones entre las distintas submuestras determinadas por las características o circunstancias pedagógicas y/o sociales que han sido consideradas relevantes y, por tanto, a tener en cuenta en nuestra investigación. Del mismo modo, hemos podido realizar estas comparaciones en función de las distintas pruebas de evaluación utilizadas en las fases pretest y posttest del estudio. La posibilidad de poder establecer comparaciones y análisis globales entre los grupos experimental y de control, al tiempo que poder realizar también contrastes en función de la pertenencia a distintos subgrupos de la muestra, de los resultados obtenidos en una determinada prueba para un subgrupo en particular, o de la interacción o convergencia entre distintos resultados obtenidos a través de distintas pruebas, incrementa la información obtenida y enriquece el análisis que puede efectuarse, al tiempo que aumenta la validez interna de la investigación.

Por otra parte, el diseño utilizado ha sido *longitudinal* si nos atenemos al hecho de que los sujetos pertenecientes a los grupos de la investigación han sido evaluados en diferentes momentos del estudio con el objetivo de valorar el efecto diferencial de la introducción del tratamiento en el grupo experimental con respecto al grupo de control.

El desarrollo de la investigación ha tenido lugar en el contexto escolar, es decir, el contexto natural donde se producen las actividades de enseñanza y aprendizaje encaminadas a la mejora de los procesos de lectura y comprensión, así como aquellas dirigidas a la evaluación de estos mismos procesos. A pesar de esto, y de que el trabajo de aplicación experimental del dispositivo de evaluación dinámica ha sido efectuado por los profesores que habitualmente atendían al alumnado participante en el estudio, la introducción de nuevas actividades y, sobre todo, la pasación de las pruebas de comprensión, inteligencia o ajuste personal social en las fases inicial y final por parte de los profesionales ajenos a los centros que han colaborado en esta tarea, lógicamente ha debido suponer un cierto elemento extraño que, por otra parte, ha debido ser compartido por todos los sujetos y grupos de la muestra, por lo que su posible influencia en los resultados habría quedado aleatorizada.

Es posible que el simple hecho de introducir un conjunto de actividades que pueden resultar novedosas, porque sitúan el énfasis del trabajo en habilidades que hasta ahora no habían sido abordadas de forma específica, aderezadas además por la puesta en práctica de una serie de patrones metodológicos que quizás resulten también novedosos, pueda explicar en parte los cambios observados en el GE de una forma artificiosa (Mora, 1998). Esta posibilidad puede quedar minimizada en parte si la duración de la experiencia supera los posibles efectos iniciales de motivación extra ante la novedad. En nuestro caso, la duración de la investigación será de un curso escolar. Aunque en nuestros planteamientos originales no hay un tiempo determinado para la aplicación de la evaluación dinámica de procesos implicados en la lectura y la posterior intervención, se estableció un periodo de seis meses para su aplicación experimental. Este tiempo parece suficiente para controlar el posible efecto al que hacíamos referencia anteriormente. Por otra parte, la posibilidad de efectos artificiosos debido a la introducción novedosa de un determinado programa también se encuentra altamente relacionada con la formación previa, la motivación y la actitud del profesorado que aplica un determinado programa de intervención o, en nuestro caso, el dispositivo de evaluación dinámica. En efecto, más allá de la bondad o calidad del dispositivo desarrollado, factores como el estilo personal del docente, su implicación y motivación en la nueva actividad o su formación en el campo específico más afín al contenido del

instrumento, pueden introducir variables difíciles de controlar con relación al peso que pueden llegar a tener en los resultados finales. Dado que, por otra parte, no era viable pensar en la aplicación experimental del dispositivo de evaluación dinámica sin contar con un grado de motivación e implicación altos, con una formación previa que hubiera contemplado al menos el contacto con algunos de los contenidos esenciales de nuestro estudio, y con una actitud y disposición positivas con relación a la formación continua y a la innovación; y dado también que la totalidad del profesorado con el que se contactó y al que se propuso su participación en el estudio, gozaba en principio de las características mencionadas más arriba (participación en proyectos y programas educativos a nivel provincial o de comunidad autónoma, desarrollo de proyectos de educación compensatoria o de atención a la diversidad, participación en Grupos de Trabajo relacionados con las necesidades educativas especiales, etc.); se optó precisamente por desarrollar el estudio contando, tanto para el grupo de tratamiento como para el grupo de control, con este conjunto de profesores y profesoras motivados e implicados ya en proyectos y programas relacionados con la mejora de la atención educativa, especialmente de sujetos con necesidades educativas especiales. El hecho de que tanto los docentes que participarían del GE como aquellos que lo harían del GC se caracterizaran por poseer un perfil similar en cuanto a los factores mencionados y considerados relevantes en el desarrollo de la investigación, parece actuar de forma suficiente como control del posible efecto diferencial de la actitud y la motivación del profesorado por la implantación de un procedimiento novedoso.

Por otra parte, y dado que la muestra principal de nuestro estudio estaba compuesta por alumnado con dificultades de aprendizaje en la comprensión, era necesario prestar una especial atención a la posibilidad de que se produjeran interacciones en la aplicación del dispositivo de evaluación dinámica, que pudieran alterar los resultados. Los alumnos que han participado en el estudio se encuentran, lógicamente, escolarizados en los distintos niveles educativos considerados en el estudio y, sin duda, están recibiendo clases de Lenguaje. Nos interesaba saber si, además de las clases de Lenguaje, se estaba trabajando algún programa educativo encaminado especialmente a la mejora de los procesos metacognitivos implicados en la lectura. La información sobre esta circunstancia la debían dar los profesores y orientadores que han participado activamente de la investigación. El objeto de tener esta información no era eliminar de la muestra a ese hipotético grupo, sino conocer esta circunstancia para poder efectuar un análisis de los datos que diera mejor cuenta de las circunstancias que pueden estar influyendo en la mayor o menor presencia de habilidades de lectura y comprensión. Por otra parte, este conocimiento sobre la posible participación de alumnos en programas de mejora de la comprensión, y, en especial, de programas de orientación metacognitiva, era especialmente importante acerca de los alumnos con dificultades de aprendizaje. En este caso, algunos de estos alumnos participan a tiempo parcial del trabajo en aulas de apoyo o logopedia; o bien forman parte de agrupamientos flexibles en los cuales se trabaja de una determinada forma, con ratios más reducidas o con la atención por parte de dos profesores en el aula; o participan de refuerzo educativo, apoyo a ciclo, etc. Estos espacios educativos suelen ser, por diversas circunstancias, el lugar donde tiene lugar en no pocas ocasiones la puesta en práctica de este tipo de programas de mejora de la comprensión. Por el motivo anteriormente explicitado, interesaba conocer el máximo de circunstancias, tanto a nivel de grupo como individualmente, que pudiera enriquecer el posterior análisis.

Las técnicas de control empleadas en nuestro estudio para minimizar posibles fuentes de variación o distorsión de los resultados han sido estas:

* En primer lugar, el establecimiento del propio diseño de investigación pretest-postest con grupo control no equivalente, facilita el control sobre las variables edad, nivel escolar, inteligencia, ajuste personal-social y rendimiento previo en lectura. La pasación de las pruebas que se recogen en el cuadro 23 y el análisis estadístico de los datos obtenidos, tanto en GE como en GC, nos ha permitido establecer el máximo grado de equivalencia posible entre ambos grupos, atendiendo a cada una de las variables especificadas. Para ello se han utilizado fundamentalmente las pruebas estadísticas de chi-cuadrado y t de Student.

La edad se ha controlado de forma que en cada uno de los grupos —GE y GC— hubiera una presencia de alumnos y alumnas equivalente en cuanto a esta variable.

El nivel escolar se ha controlado de forma que en ambos grupos existiera una presencia de alumnos y alumnas equivalente en cuanto al curso — 3º, 4º, 5º ó 6º de E. Primaria o 1º, 2º, 3º o 4º de E.S.O. — en el que se encontraban. A este respecto, y en función de los alumnos que han conformado la muestra, se ha tomado como unidad válida para establecer la equivalencia el ciclo —2º ciclo ó 3º ciclo para E. primaria y 1º o 2º ciclo para E.S.O.); es decir, si no hubiera diferencias significativas entre GE y GC en cuanto a la presencia de sujetos que cursan un mismo ciclo, se consideraría controlada esta variable.

El nivel intelectual se ha controlado efectuando la medición con la prueba estandarizada *Cattell 2-Forma A*. Se han tipificado los resultados en puntos CI con media 100 y desviación estándar 16. No debe haber diferencias significativas entre los grupos experimental y control.

El nivel de ajuste personal-social, medido con la escala *APSL*, que ha sido construida por nosotros mismos para esta investigación, también se ha controlado de forma que, tras analizar las puntuaciones pretest de los grupos experimental y control, no existieran diferencias significativas entre ellos.

El nivel de rendimiento previo en lectura ha sido controlado mediante la utilización de las pruebas *ECO 1* en E. Primaria y *ECOS* para Educación Secundaria, ambas elaboradas por nosotros para este estudio. Para considerar controlada esta variable no debían existir diferencias significativas tras el análisis de las puntuaciones obtenidas por los alumnos y alumnas pertenecientes al grupo experimental y al grupo control.

* También es importante destacar que la utilización de una batería de pruebas estandarizadas y complementarias ha contribuido a diversificar las medidas tomadas sobre los grupos experimental y control. Esto puede corregir algunas deficiencias contenidas en algunos de los instrumentos utilizados, al tiempo que permite establecer comparación entre los subgrupos (DA y NEE, por ejemplo; y también entre los subgrupos a los que un mismo profesional ha aplicado la EDPL), relativas a los resultados obtenidos en la aplicación de los distintos instrumentos por separado.

* Es necesario controlar el hecho de que los sujetos pertenecientes al grupo control reciban algún tipo de apoyo especializado durante la fase de tratamiento, ya que la aplicación de la EDPL podría considerarse así. Bastaría, en el caso del alumnado que habitualmente es atendido en las aulas de apoyo a la integración o logopedia, si reciben atención especializada por parte del profesorado de apoyo o de audición y lenguaje del centro, y son atendidos ocasional o habitualmente por los miembros del equipo de orientación o por el orientador del centro; y que, así mismo, se mantuviese esta dinámica durante el tiempo en que se aplica el tratamiento a los sujetos que participan del grupo experimental. Se trata de controlar, de este modo, la posibilidad de que las posibles ganancias obtenidas en rendimiento lector, o la mejora en la detección de las dificultades y en la intervención posterior, estén siendo comparadas con alumnado que presenta dificultades específicas de lectura y, en cambio, no esté siendo atendido. En el caso de alumnos con DA que no son atendidos en el aula de apoyo, pero que reciben apoyo a ciclo o refuerzo educativo por parte de sus tutores o de otros profesores no especialistas del Centro, estos han sido considerados a todos los efectos sujetos del GC, aunque hemos establecido posteriormente análisis comparativos en función de estas situaciones.

* Por otra parte, siendo el caso contrario al anteriormente expuesto, era de suponer que los alumnos y las alumnas que formaban parte del GC estuvieran recibiendo atención especializada adecuada a sus necesidades y, concretamente, era posible que incluso se estuvieran trabajando con ellos otros programas encaminados a mejorar su lectura. Dado que el diseño de investigación adoptado supone comparar a un grupo que recibe un tratamiento con otro que no lo recibe, y, en realidad, el GC puede estar formado por alumnado que sí estaría recibiendo otros tratamientos, se hacía necesario controlar de alguna forma este hecho. En este sentido, la diversificación de estos tratamientos, ya que los sujetos pertenecían, no sólo a centros educativos distintos, sino también a cursos distintos, hace suponer que los efectos de los tratamientos no controlados, quedarán suficientemente aleatorizados.

* Se ha controlado la presencia de alumnos y alumnas con dificultades de aprendizaje de la lectura y la de alumnado con NEE y que también presentan dificultades lectoras, en ambos grupos —GE y GC—; de forma que no haya diferencias significativas, ni en el número de sujetos de ambas tipologías presentes en cada grupo, ni tampoco en el resto de variables descritas anteriormente con respecto a ambos subgrupos.

* Se han controlado y analizado, mediante observación y entrevistas, los cambios producidos en la fase de tratamiento ajenos a la investigación y que pudieran interferir significativamente con los resultados.

Se han considerado distintos subgrupos dentro de la muestra de alumnos y alumnas con dificultades en su lectura. Estos subgrupos han permitido realizar, en función de distintas ideas previas a modo de hipótesis y de distintas variables que funcionaban como predictoras (nivel educativo, presencia de discapacidad, dificultades de aprendizaje de la lectura, privación socio-cultural, asistencia al aula de apoyo a la integración, participación de agrupamientos flexibles, participación de programas de diversificación curricular o participación de la optativa de Refuerzo de Lengua) diversas comparaciones a efectos de enriquecer los resultados, matizar las conclusiones a las que se han llegado o aportar elementos relevantes relacionados con la metodología, las condiciones de aplicación o la población beneficiaria. Se han contemplado, no obstante,

dos subgrupos principales en GE y otros dos en GC, aquellos compuestos por alumnado con dificultades de aprendizaje, en ausencia de discapacidad o deprivación socio-cultural, y aquellos compuestos por alumnado con NEE que presenta especiales dificultades en su lectura. Esto puede otorgar la posibilidad de establecer comparaciones entre los dos grupos de sujetos, tanto en el proceso de aplicación como en los resultados obtenidos. También pueden considerarse como subgrupos los integrados por los alumnos a quien se aplicará la Evaluación dinámica de procesos implicados en la lectura y la posterior intervención (en adelante EDPL) por un mismo profesional.

A continuación detallamos las comparaciones que, en este sentido, nos han interesado sobremanera. Por otra parte, en el cuadro 23, exponemos cómo se han aplicado las condiciones experimentales.

- a) Diferencias en el proceso de aplicación de la EDPL en alumnado con DA y con NEE.
- b) Diferencias en los resultados obtenidos con la aplicación de la EDPL en los distintos subgrupos contemplados: sujetos con DA, presencia de discapacidad, deprivación socio-cultural y alumnado que recibía apoyo a la integración.
- c) Diferencias en los resultados obtenidos en la batería de pruebas utilizada y en las pruebas complementarias.
- d) Diferencias en el proceso y los resultados de los subgrupos de sujetos a los que ha aplicado la EDPL un mismo profesional. Esto puede hablarnos de la puesta en práctica de elementos metodológicos diferenciadores (Mora, 1998) en los profesionales colaboradores/aplicadores, lo que, a su vez, puede contribuir a determinar elementos clave en la metodología o en los contenidos de la EDPL que contribuyan a la mejora del proceso y/o de los resultados.

4.1 Hipótesis de investigación planteadas

Las hipótesis planteadas en nuestra investigación guardan, lógicamente, estrecha relación con los objetivos expuestos anteriormente, así como con los instrumentos de evaluación empleados en el desarrollo del estudio. Las múltiples preguntas o las «certezas» que en numerosas ocasiones nos fueron asaltando durante los primeros momentos a quienes nos adentrábamos en la investigación de procesos de evaluación dinámica y en el conocimiento de los procesos de lectura, fueron tomando progresivamente forma de hipótesis de trabajo y, en cierta forma, guiando y encuadrando el desarrollo de la investigación. Planteamos a continuación cada una de estas ideas previas en relación a las cuales hemos ido desarrollando nuestro trabajo o analizando el proceso y los resultados del mismo. Una primera hipótesis experimental, que hemos considerado principal por estar relacionada directamente con los procesos de lectura y comprensión, es la relacionada con la mejora en los niveles de comprensión de textos, medidos con las pruebas de la batería criterio. Posteriormente haremos mención de otras hipótesis que, bien abarcan otros procesos contemplados en nuestra propuesta, o bien se relacionan con elementos más difícilmente validables con instrumentos de corte cuantitativo.

Cuadro 23. Condiciones Experimentales

GRUPO EXPERIMENTAL	GRUPO CONTROL
<p style="text-align: center;">Fase pretest</p> <ul style="list-style-type: none"> * <i>Test Factor "G" de Cattell, versión 2. De los denominados "libres de cultura".</i> * <i>Escala APSL (Evaluación del conocimiento y la percepción sobre los propios procesos de ajuste personal-social con relación a la lectura).</i> * <i>Prueba ECOS (Evaluación de la comprensión de textos). Esta prueba esta diseñada para E.S.O.</i> * <i>Prueba ECOI (Evaluación de la comprensión de textos). Prueba diseñada para Educación Primaria.</i> * <i>E.I.D.A.L. Cuestionario para los/as profesores/as del alumnado participante en el estudio.</i> * <i>E.I.D.A.L.-MP. Cuestionario para los padres/madres del alumnado.</i> 	<p style="text-align: center;">Fase pretest</p> <ul style="list-style-type: none"> * <i>Test Factor "G" de Cattell, versión 2. De los denominados "libres de cultura".</i> * <i>Escala APSL (Evaluación del conocimiento y la percepción sobre los propios procesos de ajuste personal-social con relación a la lectura).</i> * <i>Prueba ECOS (Evaluación de la comprensión de textos). Esta prueba esta diseñada para E.S.O.</i> * <i>Prueba ECOI (Evaluación de la comprensión de textos). Prueba diseñada para Educación Primaria.</i> * <i>E.I.D.A.L. Cuestionario para los/as profesores/as del alumnado participante en el estudio.</i> * <i>E.I.D.A.L.-MP. Cuestionario para los padres/madres del alumnado.</i>
<p style="text-align: center;">Fase de tratamiento</p> <ul style="list-style-type: none"> * <i>Evaluación dinámica de procesos implicados en la lectura (Aplicación de los instrumentos de evaluación dinámica).</i> * <i>Intervención. (Aplicación de los instrumentos de intervención y evaluación continua propuestos).</i> 	
<p style="text-align: center;">Fase de postest</p> <ul style="list-style-type: none"> * <i>Test Factor "G" de Cattell, versión 2. De los denominados "libres de cultura".</i> * <i>Escala APSL (Evaluación del conocimiento y la percepción sobre los propios procesos de ajuste personal-social con relación a la lectura).</i> * <i>Prueba ECOS (Evaluación de la comprensión de textos). Prueba diseñada para E.S.O.</i> * <i>Prueba ECOI (Evaluación de la comprensión de textos). Prueba diseñada para Educación Primaria.</i> 	<p style="text-align: center;">Fase de postest</p> <ul style="list-style-type: none"> * <i>Test Factor "G" de Cattell, versión 2. De los denominados "libres de cultura".</i> * <i>Escala APSL (Evaluación del conocimiento y la percepción sobre los propios procesos de ajuste personal-social con relación a la lectura).</i> * <i>Prueba ECOS (Evaluación de la comprensión de textos). Prueba diseñada para E.S.O.</i> * <i>Prueba ECOI (Evaluación de la comprensión de textos). Prueba diseñada para Educación Primaria.</i>

Hipótesis 1

La Evaluación Dinámica de Procesos implicados en la Lectura (EDPL), especialmente de los procesos metacognitivos, y la aplicación posterior de una intervención centrada en los procesos metacognitivos implicados en la comprensión (IPML), mejorarán el rendimiento en lectura de niños y niñas con dificultades de aprendizaje.

Actuarían como *variables independientes* el conjunto de instrumentos de evaluación dinámica de procesos implicados en la lectura, es decir, el conjunto de actividades de evaluación propuestas más los que hemos denominado finalmente patrones metodológicos de evaluación dinámica, así como la posterior intervención propuesta orientada a la solución de problemas. Como *variable dependiente*, se encontraría el grado de cambio observado y analizado en los sujetos de la investigación con relación a la hipótesis planteada. La medición del cambio en rendimiento lector se ha efectuado mediante la aplicación de pruebas de comprensión de textos. Estas pruebas, que se explicitan más adelante, se han pasado a los alumnos y alumnas participantes en la investigación, en las fases de pretest y postest del estudio, mediando entre las dos fases de la aplicación del tratamiento al alumnado participante del grupo experimental.

V.I. → Instrumentos de Evaluación Dinámica e Intervención

- *actividades de evaluación dinámica*
- *patrones metodológicos de evaluación dinámica utilizados en la mediación*
- *intervención centrada en procesos metacognitivos*

V.D. → Grado de cambio observado. Medición mediante aplicación de pruebas de comprensión de textos

Dada la clara orientación metacognitiva de nuestra propuesta, la mejora pronosticada para los sujetos a los que se aplica el dispositivo, podría tener lugar especialmente en aquellos procesos relacionados directamente con la metacognición. En este sentido, la operativización de esta medida se ve facilitada al contemplarse en las pruebas de comprensión dos dimensiones que hacen referencia a este tipo de procesos: una de ellas al *metaconocimiento estratégico* y otra a la puesta en práctica de *procesos de supervisión y autoevaluación de la propia comprensión*.

Hipótesis 2

Se esperaba que los sujetos del grupo experimental mejoraran sus niveles de ajuste personal-social como consecuencia de su participación en las sesiones de aplicación del dispositivo EDPL, aunque el cambio observado no llegara a ser estadísticamente significativo.

Aparte de la posición clave que en el estudio ocupan los procesos de lectura, nuestra investigación también ha contemplado la relevancia de la adaptación y el ajuste personal-social en el desarrollo y dominio de la lectura, y, por tanto, nos parece de gran interés considerar y analizar los posibles cambios en este ámbito que pudieran producirse en el grupo de tratamiento a causa de la participación en las actividades propuestas en el dispositivo de evaluación dinámica. Parece ser que los cambios en las actitudes y en los procesos relacionados con la motivación, los propios intereses, las expectativas o el autoconcepto, podrían tener lugar en periodos de tiempo más amplios de los que se han manejado en nuestro estudio, y a través de programas de intervención específicamente dirigidos a la modificación de actitudes concretas; por otra parte, nos parecía evidente que la aplicación de un dispositivo de evaluación dinámica, que contempla como esencial la interacción entre evaluador y sujeto y que, en contra de la posición que adopta habitualmente el evaluador, éste mantiene en nuestra propuesta una

actitud de cercanía y de apoyo constante durante el desarrollo de la actividad; podría llegar a producir una transformación significativa en la actitud, la motivación o las ideas sobre la propia capacidad de lectura y comprensión en el alumnado con dificultades de aprendizaje participante en el estudio. Por tanto, más allá de la valoración que sobre estos procesos efectuaran los profesores-aplicadores a raíz de la aplicación experimental del dispositivo, la pasación pretest y postest de la escala APSL de ajuste personal-social podría reflejar diferencias entre los grupos experimental y control una vez concluido el tratamiento. Por otra parte, aunque, como hemos apuntado, el dispositivo EDPL contiene elementos relativos a los procesos de ajuste y contempla entre sus objetivos la recogida de información que resulte valiosa al profesorado para optimizar también estos procesos en el alumno, no es un programa de intervención sobre las actitudes. De ahí que tuviéramos y expresáramos ciertas cautelas ante la posibilidad de obtener ganancias muy significativas como fruto de la aplicación de un dispositivo de evaluación dinámica de procesos implicados en la lectura.

Al igual que anotamos anteriormente, las posibles mejoras encontradas podrían tener lugar de forma preferente en aquellas dimensiones de la escala directamente relacionadas con el conocimiento metacognitivo y los procesos de autorregulación. Aunque la escala APSL nos ofrece en general información sobre procesos de *comprensión metacognitiva* (conocimiento de las propias actitudes, motivación, expectativas e intereses; percepción de la propia capacidad y de la ayuda recibida, etc.), serán las puntuaciones obtenidas en las dimensiones D8, D9 y en la dimensión que aún a las dos anteriores, D8 y 9, las que expresarían principalmente el desempeño del alumnado en el ámbito metacognitivo aplicado a la comprensión de textos y nos servirían para validar esta posibilidad.

Hipótesis 3

El grupo experimental podría incrementar sus puntuaciones postratamiento en la prueba de inteligencia general como consecuencia del trabajo desarrollado sobre los procesos metacognitivos durante la aplicación del dispositivo de evaluación dinámica.

Otro aspecto contemplado en el diseño experimental ha sido el factor de inteligencia general. Nuestro estudio no ha sido planteado con el objetivo de mostrar diferencias significativas en rendimiento cognitivo como fruto de la participación en las actividades del dispositivo; pero, por otra parte, dado que el contenido fundamental de las actividades propuestas, así como de los patrones metodológicos, era la metacognición, expresado mediante el análisis preferente de los procesos metacognitivos puestos en práctica en la actividad desarrollada en cada una de las tareas, y en la provocación consciente, por parte del evaluador, de experiencias metacognitivas en los sujetos con el objetivo de incrementar y optimizar el uso de habilidades de control y regulación conscientes de estos procesos en la comprensión de textos; cabía la posibilidad de que se produjeran cambios en el desempeño que los sujetos del grupo experimental tuvieran entre las fases inicial y final del estudio. Estos cambios podrían tener que ver con el componente metacognitivo que impregna la realización de cualquier actividad; en nuestro caso, este componente metacognitivo ha sido trabajado específicamente en el área de la lectura y la comprensión de textos, por lo que, de producirse cambios significativos a favor del grupo experimental, éstos podrían ser atribuidos adicionalmente a procesos de *transferencia* efectuados por el alumnado. Estos procesos, como pudimos comprobar en el capítulo 1, verían facilitada su aparición y puesta en

práctica mediante el trabajo sistemático sobre los procesos de reflexión metacognitiva y sobre los procesos de autorregulación.

Lógicamente, serán las puntuaciones inicial y final en la prueba de CI utilizada, las que determinen las posibles ganancias del grupo de tratamiento.

Hipótesis 4

La evaluación dinámica de procesos implicados en la lectura mejoraría el proceso de evaluación, en la medida en que ofrece al evaluador información contextualizada sobre el proceso de comprensión, las dificultades encontradas y las vías de solución.

Por otra parte, como consecuencia de la puesta en práctica de una evaluación dinámica de procesos en la lectura, y coherente con algunas de las ideas que han sido desarrolladas a lo largo del marco teórico de este trabajo, cabía la posibilidad de que el propio proceso de evaluación, en el sentido de obtener información más valiosa, contextualizada, y orientada al proceso de solución de los posibles problemas de aprendizaje descritos en el alumnado, experimentara una mejora significativa. En este sentido, el grado de satisfacción descrito por los profesionales aplicadores en las sesiones de discusión conjunta, así como en los cuestionarios, con relación a la bondad del dispositivo de evaluación e intervención para la obtención de información relevante y orientada al proceso de solución de los problemas, podrían actuar como fuentes de información a tener en cuenta para la valoración cualitativa de nuestra propuesta de evaluación. También lo haría el grado de coincidencia entre los resultados obtenidos en las pruebas de comprensión de textos y la escala de actitudes, por una parte, y la evaluación dinámica de procesos en la lectura, por otra, en cuanto a la identificación de procesos en los que existen especiales dificultades y, por tanto, en cuanto a las posibles orientaciones a la intervención.

Hipótesis 5

Los procesos metacognitivos están en la base de las dificultades de aprendizaje de la lectura. Las dimensiones metacognitivas de las pruebas de comprensión, de la escala de actitudes y del dispositivo EDPL, revelarán las mayores dificultades iniciales del alumnado en estos procesos.

En línea con lo anterior, como pudimos ver al comienzo del primer capítulo, nuestra investigación parte de una concepción de la lectura como proceso interactivo y dinámico en el que juegan un papel fundamental los procesos de control y regulación conscientes del propio proceso de lectura y aprendizaje, de los propios conocimientos y estrategias para leer y comprender. Esta constatación de la relevancia de la metacognición en el desarrollo y dominio del proceso de lectura y comprensión, nos hace pensar, a modo de hipótesis, que serían precisamente los procesos metacognitivos aquellos que estarían en la base de un número significativo de las dificultades de lectura analizadas. Será precisamente el análisis, tanto de las pruebas de comprensión de textos, como del dispositivo de evaluación dinámica, el que pueda ofrecer datos que confirmen o no lo expresado anteriormente. Dado que en las pruebas de comprensión, tanto para primaria como para secundaria, existe una dimensión de procesos de autorregulación y de metaconocimientos relacionados con la lectura, al igual que en la escala de ajuste personal-social, donde también se ha contemplado un dimensión que hace alusión al

conocimiento metacognitivo y al metaconocimiento estratégico, podemos operativizar con relativa facilidad los resultados obtenidos por el alumnado con relación a su desempeño en las cuestiones que valoran específicamente las habilidades metacognitivas. Por otra parte, en el dispositivo de evaluación dinámica, como veremos más adelante, también es posible valorar la puesta en práctica de procesos metacognitivos en la actuación de los sujetos. De esta forma, mediante la valoración y el análisis de las habilidades metacognitivas desplegadas por los alumnos durante el desarrollo de su actividad en las tareas propuestas, podremos determinar en qué medida son precisamente este tipo de habilidades las que explicarían las dificultades de aprendizaje relacionadas con la lectura encontradas mayormente en el alumnado.

Hipótesis 6

Se esperaba que el grupo que realizó la aplicación piloto obtuviera mejoras significativas en la ejecución de las pruebas de comprensión postest.

Este grupo estaba compuesto en su mayoría por alumnado que no presentaba ningún tipo de dificultades, por lo que, en función de algunos estudios ya citados en este trabajo (por ejemplo: Campione, Brown y Ferrara, 1987), así como de la experiencia acumulada por nosotros mismos, era esperable una evolución positiva del nivel de comprensión de textos, aunque no tuviera lugar una intervención específica mediante la aplicación de un programa de mejora de la comprensión lectora. Incluso en las dimensiones específicas contempladas en las pruebas, como aquellas más relacionadas con el conocimiento metacognitivo, esperábamos ganancias para este grupo.

No pensábamos lo mismo para el grupo piloto en su realización postest de la escala APSL. En efecto, el cambio esperable en el campo de las actitudes, las expectativas o la motivación relacionada con la lectura, no es comparable al que podría tener lugar en procesos más directamente relacionados con los conocimientos o los procedimientos. El hecho de que el grupo piloto estuviera formado por sujetos sin dificultades no parece en este caso ser un elemento clave para explicar un posible incremento «espontáneo» del nivel de ajuste personal-social.

Hipótesis 7

La evaluación dinámica de procesos implicados en la lectura (EDPL), como proceso de evaluación interactivo y contextualizado con relación al contenido y al proceso de evaluación, constituiría un factor explicativo del rendimiento en lectura más importante que la información obtenida de la aplicación de las distintas pruebas criterio de carácter estático utilizadas.

En este sentido, serían los análisis comparativos entre los distintos valores de predicción sobre el rendimiento escolar en lectura de las diversas puntuaciones que pueden obtenerse, los que nos ofrecerían información relevante sobre la posible contribución diferencial de ambos procedimientos de evaluación. Así mismo, estos análisis podrían mostrar la posible información adicional de las puntuaciones dinámicas (puntuaciones derivadas de la aplicación del dispositivo de evaluación dinámica, y puntuaciones de mejora) con relación a la información que es posible obtener de una evaluación de carácter convencional y estático.

4.3 Colaboradores

Un aspecto importante del estudio ha estado compuesto por el grupo de profesionales del ámbito educativo y psicoeducativo que ha colaborado en la facilitación de la puesta en práctica del dispositivo de evaluación dinámica en los distintos centros, así como en su aplicación directa al grupo experimental. Este grupo de colaboradores/aplicadores ha estado formado por maestros/as, profesores/as de apoyo, logopedas, orientadores o miembros de equipos de orientación en primaria. Esto ha sido así debido principalmente a que la muestra ha estado compuesta esencialmente de sujetos «aislados», es decir, no componían, salvo los integrantes de la muestra piloto, un grupo-clase. Se trataba de alumnos y alumnas que presentaban dificultades de aprendizaje de la lectura y, normalmente, no se encuentran todos estos alumnos en una misma clase. Por este motivo, cobraba especial importancia la colaboración y la motivación de profesionales que, en distintos centros educativos, facilitarían la aplicación a algunos sujetos —escogidos en función de pertenecer al grupo de alumnos y alumnas con dificultades de aprendizaje de la lectura y cuyas características quedarán especificadas más adelante—, de las pruebas previstas para poder establecer comparaciones en las distintas fases del estudio, así como de las pruebas complementarias para el profesorado tutor de los alumnos y para las familias; así mismo, algunos de estos profesionales han aplicado directamente el dispositivo de evaluación dinámica de procesos implicados en la lectura en la fase de tratamiento del estudio al grupo experimental.

Se ha contado con 12 de estos profesionales que han colaborado en la puesta en práctica del dispositivo de evaluación dinámica EDPL en la fase de tratamiento de la investigación. El hecho de contar con estos doce aplicadores del dispositivo nos permitía observar y analizar distintas formas de llevar a la práctica el instrumento, y valorar aquellos patrones metodológicos de evaluación dinámica que resultaban más efectivos con relación a la valoración del proceso que efectuaban los mismos aplicadores, así como con relación a los resultados obtenidos. Se comenzaría a realizar esta valoración en sesiones conjuntas de discusión, en las que participaría el grupo de aplicadores, así como mediante la realización de un cuestionario que trataría de analizar las dificultades encontradas, los elementos que podrían mejorarse, los elementos de los instrumentos de evaluación considerados *clave*, es decir, que permiten obtener una información más valiosa para la posterior intervención, así como los elementos clave del proceso *evaluación dinámica - intervención - evaluación continua*, que permiten explicar la posible mejora en los procesos de lectura.

Este grupo de aplicadores estaba formado por 4 profesores/as de apoyo a la integración, una profesora tutora con formación en Educación Especial y dos integrantes de equipos de orientación educativa (EOE), todos en el ámbito de actuación de educación primaria; así mismo, estaba integrado por cuatro profesoras de apoyo a la integración de educación secundaria y una orientadora de un IES. La formación inicial y continua de estos aplicadores se caracterizaba por su especialización en la atención educativa a las necesidades educativas especiales, contando entre ellos con ocho profesionales que habían realizado la licenciatura en psicología o pedagogía.

También se ha contado con otros profesionales, que pertenecían igualmente al ámbito educativo y psicoeducativo, formado igualmente por maestros/as, profesores/as de apoyo a la integración, logopedas y orientadores/as, y que, junto con el grupo de alumnos que se especificará más adelante, han formado parte del GC. Su labor, en este

caso, ha sido facilitar la pasación de las pruebas de la batería a los alumnos y alumnas en las fases pretest y postest del estudio, así como las pruebas complementarias durante el transcurso del mismo. Este grupo de profesores que ha prestado su colaboración con aquellos grupos de alumnos que conformaban el GC, estaba compuesto por dos profesoras de apoyo a la integración y ocho profesores/as tutores/as, todos ellos pertenecientes al ámbito de actuación de educación primaria; así como por un profesor tutor de Lengua y dos orientadores/as de dos IES, pertenecientes al ámbito de actuación de educación secundaria. Es preciso tener en cuenta que muchos de los sujetos que conformaron el GC se encontraban escolarizados en los mismos centros que albergaban los grupos de alumnos que formaban parte del GE. En estos casos, los alumnos del GC no estaban siendo atendidos por el profesorado que aplicó el dispositivo de evaluación dinámica en la fase de tratamiento; se trataba de alumnado que presentaba, como veremos posteriormente más detalladamente, dificultades de aprendizaje relacionadas con la comprensión, pero que no llegaba a formar parte del grupo de alumnos que recibía apoyo a la integración, recibiendo algún tipo de apoyo o refuerzo educativo, o bien participando de medidas de adaptación grupal como los agrupamientos flexibles. En cualquier caso, la implicación y la motivación de este grupo de profesores/as colaboradores han sido excelentes, facilitando en todo momento la aplicación de las pruebas pretest y postest, al tiempo que la cumplimentación de los cuestionarios o las plantillas de valoración del rendimiento escolar en el área de Lengua por parte de sus compañeros, cuando esto fue necesario.

También se ha contado con varios colaboradores para la aplicación de las pruebas estandarizadas en las fases pretest y postest del estudio. Estos colaboradores han sido finalmente seleccionados de entre el grupo de alumnos que había finalizado la diplomatura de Magisterio o estaba finalizando la licenciatura de Psicopedagogía y que mantenían contacto con el departamento de Psicología Evolutiva y de la Educación de la Universidad de Sevilla, bien por estar participando de programas de alumnado interno o bien por haberlo hecho recientemente. Lógicamente, también fueron seleccionados en función de su disponibilidad para el trabajo a realizar. Este grupo de colaboradores estuvo compuesto en la fase pretest por tres diplomadas en Magisterio, una de las cuales estaba finalizando sus estudios de Psicopedagogía. En la fase postest, como explicaremos más adelante, debido a problemas de índole personal, tan solo una de estas colaboradoras pudo seguir con su labor y hubieron de seleccionarse tres compañeros/as más para completar la fase final de aplicación de las pruebas. Estos tres nuevos colaboradores que se unieron a la que ya completó la aplicación en la fase pretest, habían sido seleccionados de igual forma que las primeras y su formación era similar: dos estudiantes de último curso de Psicopedagogía y un estudiante de Magisterio.

Para la validación de expertos de las pruebas elaboradas por nosotros para esta investigación, se ha contado con profesionales de la educación del ámbito académico (orientadores/as y profesores/as), todos ellos con experiencia y formación en Psicología educativa o Pedagogía. En total han sido cuatro los profesionales que han validado las pruebas, repartiendo su trabajo en función del ámbito de actuación en el que se desarrolla su trabajo. De esta forma, los cuatro han valorado la escala de actitudes APSL, por tratarse de una prueba que, básicamente, ha sido aplicada de igual forma en primaria y en secundaria; tres de ellos han realizado su valoración sobre la prueba de comprensión de textos ECO 1, y una orientadora ha valorado la adecuación de la prueba ECOS para la evaluación de la comprensión de textos en educación secundaria.

En esquema, la participación de los distintos colaboradores en el estudio queda concretada de la siguiente forma:

* La colaboración de los 12 profesores/orientadores ha contemplado las siguientes tareas:

- Facilitar la pasación de las pruebas pretest y postest al alumnado participante en el estudio.
- La aplicación del dispositivo EDPL con el GE.
- Cumplimentar el Cuaderno de hojas de registro y evaluación, en el que debe plasmarse la valoración sobre la mediación efectuada y sobre el proceso de aplicación de cada una de las actividades de evaluación dinámica realizadas.
- Análisis en el *Grupo de Trabajo* del CEP del proceso de aplicación del dispositivo y de los resultados obtenidos.
- Cumplimentar los Cuestionarios sobre Expectativas e Ideas relacionadas con las dificultades de aprendizaje (EIDAL) y facilitar a los profesores tutores del alumnado participante en el estudio estos mismos cuestionarios, así como a las familias los Cuestionarios sobre Expectativas e Ideas relacionadas con las dificultades de aprendizaje (EIDAL-MP), versión madres/padres.
- Facilitar a los profesores tutores del alumnado participante la plantilla de valoración sobre el rendimiento escolar en el área de Lengua, así como cumplimentar esta misma plantilla de valoración.

* La participación de los otros 13 profesores colaboradores, especialmente de aquellos que han colaborado desde centros en los que solo había alumnado que formaba parte del GC, ha significado:

- Facilitar la pasación de las pruebas pretest y postest al alumnado participante en el estudio.
- Cumplimentar los Cuestionarios sobre Expectativas e Ideas relacionadas con las dificultades de aprendizaje (EIDAL) y facilitar a los profesores tutores del alumnado participante en el estudio estos mismos cuestionarios, así como a las familias los Cuestionarios sobre Expectativas e Ideas relacionadas con las dificultades de aprendizaje (EIDAL-MP), versión madres/padres.
- Facilitar a los profesores tutores del alumnado participante la plantilla de valoración sobre el rendimiento escolar en el área de Lengua, así como cumplimentar esta misma plantilla de valoración.

* La participación de los colaboradores que han aplicado las pruebas pretest y postest en los centros, se ha concretado en las siguientes tareas:

- Participar de tres/cuatro sesiones de formación inicial con el objetivo de conocer en profundidad cada una de las pruebas en cuestión, así como de tener la oportunidad de poder practicar su uso de forma experimental, analizando las dificultades que podía conllevar su aplicación posterior con los grupos del estudio.
- Mantener reuniones semanales durante el periodo de tiempo que durara la aplicación pretest y postest, con el objetivo de hacer entrega de las pruebas

finalizadas, así como de revisar la aplicación que se estaba realizando para poder corregir posibles desviaciones de los criterios acordados en la formación inicial.

Formación de los colaboradores/aplicadores del dispositivo EDPL

La puesta en práctica de la Evaluación Dinámica de Procesos implicados en la Lectura (EDPL) ha requerido de un proceso sistemático de formación. Los 12 profesionales participantes fueron, previamente a su utilización, instruidos en el conocimiento y en el uso práctico de los instrumentos de evaluación-intervención. Esta formación se efectuó mediante la participación en sesiones de trabajo con la persona responsable del proyecto. Las sesiones de formación se efectuaron en el contexto de la participación en un *Grupo de Trabajo* del C.E.P. de Alcalá de Guadaíra y ocuparon las primeras cuatro sesiones del trabajo desarrollado en el Grupo. Posteriormente tuvieron lugar sesiones de seguimiento y evaluación del proceso de aplicación del dispositivo durante la fase de aplicación del mismo y sesiones de evaluación de su aplicación y de los resultados obtenidos, una vez finalizada su aplicación. En el capítulo en el que queda recogido el *Procedimiento* seguido en nuestra investigación, se explicitan algunos de los elementos conceptuales, materiales y metodológicos, que integraron el contenido de las sesiones de formación de los aplicadores.

5. Participantes

5.1 Composición y distribución de la muestra

El proceso de selección de los sujetos candidatos a participar en este estudio se ha realizado mediante el contacto con escuelas públicas de nuestro entorno (Sevilla y su provincia, y Puerto Real, en Cádiz) y en función de las posibilidades y las facilidades otorgadas por éstas para este trabajo. Los participantes, por tanto, han sido algunos miembros de la comunidad educativa de estos centros: alumnos y alumnas con dificultades de aprendizaje en la lectura, profesorado que atiende habitualmente a estos niños y niñas (normalmente tutores, profesores de apoyo o logopedas), miembros de los Equipos de Orientación Educativa (E.O.E.), orientadores y orientadoras de E.S.O. y padres y madres del alumnado seleccionado. Aunque los sujetos experimentales de la investigación han sido los alumnos y las alumnas con dificultades en lectura, a través de los cuales se ha pretendido establecer los posibles cambios en la capacidad de comprensión como consecuencia de la introducción de las diversas variables, es importante reseñar la relevancia que ha cobrado en nuestro estudio la participación de los otros miembros de la comunidad educativa, entre los cuales se encuentran todos los mencionados anteriormente. De ahí que algunos de los instrumentos que se hayan aplicado a lo largo del estudio, hayan estado dirigidos específicamente al profesorado o a los padres y madres.

Aunque en nuestro Proyecto de Investigación se recogía la participación aproximada de una centena de alumnos y alumnas, la ampliación posterior, en función del interés que su inclusión podía aportar al estudio, a un segundo grupo control compuesto por alumnado que no presentaba problemas de aprendizaje, ha supuesto en la práctica que la muestra definitiva haya estado compuesta por 335 estudiantes de educación primaria y secundaria. La muestra se ha distribuido de la siguiente forma: el GE de primaria ha estado conformado por 32 alumnos, mientras que el GC ha contado con 36 sujetos; en secundaria, el GE lo formaban 28 alumnos y el GC un total de 37. En total han sido 133 los alumnos con DA en la lectura que han participado del estudio, de los cuales 60 lo han hecho formando parte del GE, en el cual se ha implementado el tratamiento, es decir, los instrumentos de evaluación dinámica de procesos implicados en la lectura; y 73 alumnos han participado del GC, el cual ha compartido todas las condiciones experimentales del GE, excepto la aplicación del tratamiento. En cuanto al resto de alumnos y alumnas, es decir, aquellos que han formado parte de ese segundo grupo control, compuesto por sujetos sin dificultades especiales en lectura, su número ha sido de 202.

Como explicaremos posteriormente de forma más detallada, tras la aplicación piloto de las pruebas de comprensión diseñadas y utilizadas en las fases pretest y postest, y una vez que estas pruebas no sufrieron cambios significativos desde esta aplicación, se consideró oportuno recoger la información sobre los resultados obtenidos de su pasación en los distintos grupos, volviendo a realizar las pruebas en la fase postest al tiempo que se efectuaba con los grupos de alumnos con dificultades. El objetivo esencial de la inclusión de este segundo Grupo Control (GC2) fue el de ofrecer una medida del cambio en el rendimiento en comprensión sin la intervención específica que

suponía la aplicación del procedimiento de evaluación dinámica EDPL y con alumnado sin dificultades de aprendizaje. Se esperaba que el alumnado que no presentaba dificultades experimentara un crecimiento «espontáneo» en cuanto a su capacidad de comprensión de textos, incrementando sus puntuaciones en el postest con respecto a las obtenidas en la fase pretest. Este crecimiento espontáneo, en el sentido de que no media la aplicación experimental que sí se otorga a otro de los grupos que conforman el estudio, ha sido aludido por numerosos investigadores con relación al incremento que se produce, a nivel cuantitativo y cualitativo, en la adquisición y uso de conocimientos y estrategias en el alumnado sin dificultades (Campione, Brown y Ferrara, 1987; Martín y Marchesi, 1990; Martín, 1999; Navarro y Mora, 2002) en comparación con el grupo de alumnos y alumnas que presentan problemas de aprendizaje.

Han participado un total de 13 centros educativos, de los cuales 12 eran públicos y 1 lo era concertado; por otra parte, 7 centros lo eran de educación primaria y 6 de educación secundaria. La distribución del alumnado participante por centros, así como su grupo de participación queda reflejada en la siguiente tabla:

Tabla 1. Distribución de sujetos por centros y grupos de participación

Centros	Nº de alumnos	GE	GC1 (con DA)	GC2 (sin DA)
CEIP Lope de Rueda (Sevilla)	3	3	0	0
C. Concertado Sta. Isabel (Sevilla)	4	0	4	0
CEIP El Ruedo (Arahal)	6	5	1	0
CEIP Andalucía (Sevilla)	22	10	3	9
CEIP M. Sánchez Alonso (Arahal)	183	8	25	150
CEIP Antonio Machado (Morón de la Fra.)	6	6	0	0
CEIP Vicente Neria (Coria del Río)	3	0	3	0
IES Torreblanca (Sevilla)	10	10	0	0
IES Los Azahares (La Rinconada)	10	0	10	0
IES Mariana Pineda (Montequinto)	5	5	0	0
IES Leonor de Guzmán (Alcalá de Gdra.)	58	13	2	43
IES la Jarcia (Puerto Real)	22	0	22	0
IES Torre de los Guzmanes (La Algaba)	3	0	3	0
Totales	335	60	73	202

En cuanto a la distribución por cursos, la mostramos igualmente en la siguiente tabla junto con el grupo de participación. Puede comprobarse que el número de alumnos pertenecientes a tercer curso de primaria es muy bajo; la causa hay que hallarla en el hecho de que, en la práctica, la mayoría de los profesores colaboradores han observado grandes dificultades en la aplicación de las pruebas de evaluación establecidas a niños y

niñas por debajo de cuarto de primaria, máxime cuando se trataba de alumnado que presentaba dificultades en su aprendizaje.

Tabla 2. Distribución del alumnado por cursos y grupos de participación

Curso o nivel educativo	Nº de alumnos	GE	GC1 (con DA)	GC2 (sin DA)
3º primaria	3	1	2	0
4º primaria	30	4	10	16
5º primaria	109	18	15	76
6º primaria	85	9	9	67
1º secundaria	17	6	11	0
2º secundaria	48	7	16	25
3º secundaria	13	5	8	0
4º secundaria	30	10	2	18
Totales	335	60	73	202

5.2 Dificultades en la composición final de la muestra. Las características de los sujetos participantes y las decisiones adoptadas

Por otra parte, como era de esperar en un estudio de estas características, que se ha desarrollado en el marco escolar y durante todo un curso académico, se ha producido un cierto número de bajas entre el alumnado participante por diversos motivos. En primer lugar, 9 alumnos que iban a formar parte del GE no han podido finalmente conformar este grupo debido básicamente a cuatro tipos de problemas: absentismo prolongado, cambios de centro, abandono escolar y no haber asistido al centro cuando se han realizado las pruebas posttest; en este sentido, se optó por eliminar de forma general a todos aquellos alumnos a los que le faltaran 3 o más pruebas, ya que, en realidad eso implicaba la imposibilidad de poder establecer algún tipo de comparación. También en el GC1 otros 9 alumnos se han quedado fuera de la muestra definitiva. Las causas han sido similares a las expuestas para el GE. Igualmente, para el GC2, es decir, para el grupo de alumnos y alumnas sin DA que han formado parte de un segundo grupo de control, ha existido un cierto nivel de mortandad experimental, que, en cualquier caso, como ha ocurrido con los grupos anteriores, no parece haber representado una influencia significativa para los resultados del estudio. En algunos casos, el abandono de la muestra experimental de la investigación, ha tenido lugar cuando ya se habían realizado algunas de las pruebas pretest; en estos casos, si el alumno formaba parte del GC2 y, por tanto, había recibido la aplicación piloto, se ha optado por incluirlo en los resultados grupales aunque, finalmente, no pudiera realizar las pruebas posttest. En este caso, se ha valorado la comparación pretest-posttest grupal y la ausencia de unos pocos alumnos en el posttest no tendría ningún tipo de consecuencias para el establecimiento de conclusiones. En cambio, en los GE y CG1, las ausencias de alumnos en la realización de pruebas pretest o posttest tenía una mayor relevancia. Dado que algunos de estos alumnos han participado con normalidad en la aplicación experimental del dispositivo, pero se han ausentado en algunas de las pasaciones de pruebas, no hemos

querido eliminar de la muestra a estos alumnos a no ser, como decíamos anteriormente, que le faltaran tres o más de las pruebas, en cuyo caso no sería posible, ni significativa, ningún tipo de comparación. De otro modo, cuando se ha producido alguna ausencia, pero se tienen medidas pretest y posttest del resto de pruebas, se ha optado por incluir a estos alumnos en los resultados grupales de las pruebas realizadas, por lo que es posible que algunas comparaciones que efectuemos contengan un mayor número de efectivos que otras.

Uno de los problemas a los que se hacía referencia anteriormente, el absentismo, ha tenido su influencia más allá de la realización de las pruebas pretest y posttest; se ha manifestado sobre todo para el GE en el transcurso de la aplicación del dispositivo. El grado de absentismo ha sido significativo especialmente para algunos alumnos pertenecientes a dos centros: el IES Torreblanca y el CEIP Andalucía, ambos situados en Z.A.E.P. (zonas de actuación educativa preferente) de Sevilla. Aunque se ha contado en la muestra con varios grupos pertenecientes a centros situados en zonas deprivadas, con niveles superiores de riesgo de absentismo escolar o situaciones de deprivación social y cultural, podríamos catalogar la totalidad de la muestra, tanto en educación primaria como en secundaria, como perteneciente a un nivel socio-cultural medio-bajo.

El hecho de haber incluido en la muestra a grupos de alumnos y alumnas escolarizados en centros situados en zonas deprivadas, ha respondido, por una parte, a que en estos centros se encontraban profesores u orientadores que estaban dispuestos a participar en la investigación; por otra, a la relevancia de establecer comparaciones en cuanto a las medidas efectuadas en la investigación, en función de situaciones de deprivación socio-cultural. Como exponíamos en otro estudio (Navarro y Mora, 2002), la población escolar en estas situaciones ha sido la principal destinataria de muchas de las investigaciones sobre la aplicación dinámica de sistemas de evaluación (Feuerstein, 1991; Mora, 1998), y, aunque nuestro estudio no se dirige a la valoración de funciones cognitivas generales como la gran mayoría de los estudios referidos y se centra de forma especial en la evaluación de un dominio específico como la lectura, la posibilidad de contar con un grupo de alumnos que respondiera de una manera clara a estas características, era motivo más que suficiente para su inclusión en la muestra. Por otra parte, conocíamos que en uno de los centros mencionados se estaban llevando a cabo experiencias de apoyo basadas en agrupamientos flexibles para contenidos especialmente relacionados con el aprendizaje y desarrollo de la lectura, lo cual constituía un motivo extra para contar con una muestra de estos alumnos, ya que podía ser interesante averiguar en qué medida estos aspectos organizacionales y de estructuración de tiempo y contenido en función de los niveles de competencia curricular de los alumnos, se reflejarían en los resultados obtenidos.

La catalogación del alumnado participante como alumno o alumna con dificultades de aprendizaje de la lectura, deprivación socio-cultural o discapacidad asociada a problemas en la comprensión de textos, ha sido realizada desde una doble vertiente; por una parte, ha tenido lugar una designación previa por parte del profesorado colaborador participante en el estudio (profesores de apoyo a la integración, orientadores y logopedas) (García Sánchez y Marbán, 2003). Como ya comentábamos en otro lugar (Navarro y Mora, 2002), esta práctica tiene ciertas ventajas, como por ejemplo el hecho de que la catalogación de alumno con dificultades en la lectura venga dada por aquellos profesionales especializados que desarrollan cotidianamente su trabajo con este grupo de alumnos y alumnas y mejor conocen su nivel de competencia. Una ventaja adicional

y especialmente relevante de esta opción la constituye el hecho de que la designación previa de los profesores y orientadores sobre las dificultades de aprendizaje, podía convertirse en un criterio de validez externa posterior en función de los resultados obtenidos en las pruebas. En este caso, el riesgo que podría suponer el hecho de que no se tuviera un conocimiento detallado sobre el nivel de competencia en lenguaje escrito por parte de los profesores, parece estar minimizado dado que la muestra está mayoritariamente compuesta por alumnado que participa del trabajo desarrollado en las aulas de apoyo a la integración y de audición y lenguaje, siendo habitualmente bien conocido por el profesorado especialista y por los orientadores de los centros. En otros casos, en los que el alumnado con dificultades no recibía apoyo por parte del profesorado de educación especial, lógicamente ha sido más difícil determinar si se cumplían las características esenciales para formar parte de la muestra; en ocasiones ha sido necesario contar con la colaboración del profesorado tutor del alumno, que nos ha informado sobre el alumno en cuestión. Este ha sido el caso de algunos sujetos que finalmente han formado parte del GC1, los cuales han sido seleccionados con base en la información proporcionada por los tutores al profesorado de apoyo o a los orientadores que han colaborado en el estudio.

Una vez comenzado el estudio, se entregó a los profesores colaboradores una plantilla (Anexo V) en la que figuraban los nombres de los alumnos y alumnas participantes y una serie de rejillas en las que debían colocar si el alumno manifestaba dificultades de aprendizaje de la lectura (DA), discapacidad, desventaja socio-cultural, si recibía algún tipo de apoyo (apoyo a la integración, refuerzo educativo, programas de agrupamientos flexibles, etc.), si participaba de programas de diversificación curricular (DC), o si lo hacía de las clases optativas de Refuerzo de Lengua. Con la cumplimentación de esta plantilla se obtenía una información detallada por parte del profesorado y los orientadores sobre la totalidad de la muestra con dificultades de lectura.

Por otra parte, la catalogación de alumnado con dificultades de aprendizaje de la lectura podía venir dada también en función de los resultados obtenidos en las pruebas pretest del estudio, sobre todo en las pruebas de comprensión y en Cattell. De este modo podríamos establecer hasta tres criterios para establecer las dificultades de aprendizaje: (a) valoración del profesorado/orientadores; (b) puntuación en la prueba de comprensión; y (c) puntuación en Cattell. En este sentido, era posible que un alumno cumpliera un solo criterio, o dos, o los tres. Se podía, de este modo, analizar y reanalizar los resultados en función de haber tomado uno, dos o los tres criterios. Esto cobraba especial importancia para el grupo de alumnos y alumnas que formaron parte del GC1 y que no recibían atención educativa especial. Por ejemplo, un alumno podía haber sido catalogado con DA por el profesor, pero quizás la realización de las pruebas de comprensión revelara un resultado dentro de la media del grupo de alumnos que no presentaban dificultades, e igualmente con respecto a sus resultados en la prueba de inteligencia general; en ese caso, lógicamente, era preciso tener una especial cautela con las conclusiones que podíamos extraer de los análisis comparativos efectuados, si finalmente se incluía a este sujeto en el grupo de alumnos con dificultades de aprendizaje.

Las características generales más relevantes del grupo de sujetos con dificultades en el aprendizaje seleccionados, han sido finalmente:

- a) *pertenecer al grupo de niños y niñas con dificultades en el aprendizaje de la lectura, ya se manifiesten estas dificultades en el reconocimiento de palabras, en la comprensión global de los textos, o en ambas dimensiones;*
- b) *estar matriculados en los cursos 3º, 4º, 5º ó 6º de primaria, o en cualquiera de los cuatro niveles de educación secundaria obligatoria.*

Con respecto a la primera de las características definidas y, dentro de ésta, al grupo de niños y niñas que presentan dificultades en el reconocimiento de palabras, es necesario destacar que debían cumplir ciertas condiciones. Emilio Sánchez (1999) recoge las siguientes:

- que la capacidad intelectual de los sujetos sea normal (por ejemplo un CI no inferior a 85);
- que se aprecie un retraso de al menos dos años entre la capacidad general (evaluada a través del CI) y el rendimiento en la lectura (evaluada a través de una prueba de comprensión de textos);
- que haya contado con la oportunidad de aprender, esto es, que haya recibido una enseñanza convencional: asistencia regular a las aulas con una formación adecuada;
- que no haya una causa que por sí misma pueda explicar el retraso, sean problemas sensoriales importantes, problemas emocionales que hayan impedido la participación del alumno en las experiencias de aprendizaje y enseñanza, o discapacidad psíquica.

El grupo de niños con dificultades en la comprensión global de los textos, así como aquellos que manifiestan problemas en ambas dimensiones, debería cumplir estas mismas condiciones, aunque, dado que este tipo de dificultades posee un grado de especificidad mucho menor, cabía esperar que los problemas manifiestos de lectura se encontraran asociados a otros problemas más generalizados. Entre éstos podemos destacar junto a Alonso Tapia, Carriedo y Mateos (1992):

- escasa o pobre utilización de la capacidad de supervisión y evaluación de la propia comprensión
- escaso conocimiento de estrategias adecuadas para afrontar tareas de lectura y comprensión
- escasa capacidad en el manejo y uso de estrategias de lectura

También podríamos encontrarnos con una escasa o pobre utilización de la memoria de trabajo, escaso control y regulación de la atención y en general de todos los procesos cognitivos necesarios para acceder a la comprensión, dificultades para relacionar o conectar ideas entre sí, etc.

En general, podríamos operativizar este grupo de forma que fueran incluidos todos aquellos alumnos que presentaran problemas de aprendizaje de la lectura y no pudieran ser incluidos en el primer grupo. En este sentido, dado que, por otra parte, muchos de los niños que presentan dificultades de aprendizaje en la lectura, y con los cuales se trabaja habitualmente en las tutorías y en las aulas de apoyo a la integración, son alumnos y alumnas que presentan también otras necesidades educativas especiales asociadas a discapacidad, nosotros incluiremos en nuestro estudio a estos sujetos, dentro de este segundo grupo mencionado, procediendo posteriormente, en el análisis de los

resultados obtenidos con la aplicación de la evaluación dinámica de los procesos de lectura, a una valoración diferencial de los posibles efectos de la aplicación del tratamiento en ambos grupos (DA y NEE), y, valorando igualmente, las variables y los elementos, tanto metodológicos como de contenido mismo de los instrumentos, que habrían influido en los resultados, aparte de la discapacidad.

Con respecto a la segunda característica operativizada, queremos destacar que optamos finalmente por el curso escolar como criterio esencial de participación en vez de la edad, ya que, sobre todo en cuarto curso de educación secundaria hay un número importante de alumnado que presenta DA o discapacidad que supera los 15 o los 16 años. Dadas las características de nuestra investigación, nos ha parecido oportuna la inclusión de todos los alumnos que estuvieran en los cursos citados. En cualquier caso, como es lógico, la gran mayoría del alumnado participante estaría entre los 8-9 años y los 15-16.

Un aspecto que ha ocasionado algunas dudas en el profesorado que ha colaborado con nosotros en la aplicación del dispositivo ha sido la catalogación de alumnos con deprivación socio-cultural. Una parte importante de estos alumnos lo componen en nuestra muestra alumnos de un centro situado en una zona de actuación educativa preferente, por lo que la determinación por parte del profesorado de que se trataba de alumnos en situación de deprivación social no dejó lugar a muchas dudas. En los demás casos, los profesores y los orientadores han debido determinar si un alumno se encontraba en esta situación de riesgo, aún cuando la zona en la que estuviera enclavado el centro no fuera especialmente deprivada. A la hora de aclarar y unificar los criterios para que un alumno fuera considerado en nuestra investigación como alumno con deprivación socio-cultural, se han comentado con los profesores las características que habíamos considerado; básicamente se trataba de alumnos de los que se tuviera conocimiento de que vivían bajo condiciones sociales y familiares que no facilitaban, antes bien, perjudicaban, su desarrollo social, cultural y escolar.

En la siguiente tabla podemos observar la distribución de los sujetos en función de la catalogación efectuada por el profesorado participante en el estudio, así como su pertenencia a los grupos experimental o control. Hemos tomado como referencia para los análisis de composición de la muestra el grupo de alumnos y alumnas con dificultades de aprendizaje que ha formado parte de los grupos experimental y control (GC1). Posteriormente, cuando abordemos la aplicación piloto de las pruebas de comprensión y actitudes, haremos referencia más detallada al segundo de los grupos de control conformados. Como podrá comprobarse por los números totales, a algunos sujetos le fueron asignadas más de una categoría; esto es así debido a que algunos de estos grupos no se han tratado, ni fueron propuestos al profesorado que debía encuadrar en ellos al alumnado al que atiende, como excluyentes entre sí. De este modo, alumnos que presentaban discapacidad también aparecían en las plantillas como alumnado con dificultades de aprendizaje de la lectura, ya que, aunque en el marco de nuestro estudio, podía quedar incluido dentro de la especificación de alumno con discapacidad ligada a problemas en la comprensión, parecía más visible y fácil de manejar la opción de señalar ambas casillas. Por otra parte, también resultaba lógico y esperable que alumnos con DA, discapacidad o deprivación socio-cultural recibieran algún tipo de apoyo o participaran en la optativa de Refuerzo de Lengua, en secundaria. En la plantilla que fue ofrecida a los colaboradores, figuraban una serie de indicaciones que pretendían facilitar la identificación de las distintas categorías y la unificación de criterios. Las abreviaturas

utilizadas en la tabla son estas: DA (dificultades de aprendizaje de la lectura); Discap. (presencia de discapacidad ligada a problemas en la lectura); DS (deprivación socio-cultural); Recibe apoyo: AI (apoyo a la integración y/o audición y lenguaje), AC (apoyo a ciclo o refuerzo educativo), AF (agrupamientos flexibles); DC (participa de programas de diversificación curricular); RL (participa de la asignatura Refuerzo de lengua).

Tabla 3. Asignación de los sujetos de la muestra por categorías de dificultad

Curso	DA		Discap.		DS		Recibe apoyo						DC		RL	
	GE	GC	GE	GC	GE	GC	AI		AC		AF		GE	GC	GE	GC
							GE	GC	GE	GC	GE	GC				
E. primaria	32	38	7	4	16	11	12	4	2	20	10	3				
E. secundaria	28	35	6	12	4	3	19	15					9		14	20
<i>Totales</i>	<i>60</i>	<i>73</i>	<i>13</i>	<i>16</i>	<i>20</i>	<i>14</i>	<i>31</i>	<i>19</i>	<i>2</i>	<i>20</i>	<i>10</i>	<i>3</i>	<i>9</i>	<i>0</i>	<i>14</i>	<i>20</i>

Por otra parte, en una investigación como la que presentamos, encuadrada dentro de la investigación educativa, tratando de extraer información relevante para la comprensión de procesos cognitivos y también educativos dentro del contexto escolar y con instrumentos de medida lo más parecidos posible a tareas que los alumnos tengan la oportunidad de realizar; algunos requisitos de los estudios experimentales no podían ser cumplidos totalmente. La selección de la muestra ha venido dada por la condición de presentar dificultades en el aprendizaje de la lectura, por lo que, de este modo, por ejemplo, la aleatorización de los sujetos no podía ser completa. Es más, en nuestro caso, esta selección en función de que se cumplieran determinadas condiciones, ha sido la que ha permitido poder llevar a cabo el estudio, ya que se trataba de valorar la aplicación dinámica de un dispositivo de evaluación en alumnado con dificultades de aprendizaje. Igualmente, la equivalencia total entre los grupos que posteriormente podían ser comparados, no podía ser total. Dado que la muestra ha sido seleccionada en función de la presencia de dificultades en la lectura, y dado también que el hecho de que se haya aplicado el dispositivo a un grupo determinado y no a otro ha respondido la mayoría de las veces a la facilidad que han otorgado algunos profesores, así como al mayor interés mostrado por éstos y por algunos orientadores de los centros; cabía la posibilidad de que existiera un efecto no controlado debido al quehacer de unos profesores con respecto a otros en el tratamiento cotidiano anterior de las estrategias de comprensión en sus clases. Sin embargo, la relativamente amplia muestra recogida y la diversidad de aulas, profesores y centros educativos participantes de la investigación, así como la mencionada implicación educativa del profesorado que participó, tanto del GE como del GC; han debido funcionar suficientemente como elementos que aleatorizaran cualquier efecto que se pudiera haber producido en el sentido mencionado. Por otra parte, la información recogida por los profesores y orientadores que han colaborado más estrechamente en el estudio, no ha revelado que hubiera tenido lugar especialmente la puesta en práctica de programas específicos encaminados a la mejora de procesos de comprensión, ni en el GE ni en el GC.

5.3 Análisis de la equivalencia entre los grupos

El análisis de la equivalencia entre los grupos control y experimental contemplará diversos factores: el sexo de los sujetos, la edad y el nivel educativo, el nivel intelectual,

el rendimiento en lectura y el nivel de ajuste personal-social. Nos detendremos a continuación en cada uno de ellos para analizar la presencia o ausencia de diferencias de partida entre los grupos de tratamiento y de control. Este análisis nos será posteriormente de utilidad, una vez que nos adentremos en la valoración e interpretación del proceso y de los resultados obtenidos en el estudio.

Equivalencia con relación al sexo

Con respecto a la equivalencia de los sexos en la muestra, se ha descrito en este sentido que pueden existir diferencias en cuanto a rendimiento y pautas de desarrollo en algunas áreas o ámbitos de estudio. Entre estas áreas se encuentra el lenguaje, en el que los resultados de algunas observaciones parecen decantarse a favor de un mayor rendimiento por parte de las niñas (Valmaseda, 1999; del Río, 1999). Igualmente, se ha hecho mención en la literatura sobre una mayor precocidad de las niñas en la adquisición de habilidades lingüísticas y de adaptación social (Mora, 1998; Freixas y Luque, 1998), al tiempo que también se ha descrito una mayor prevalencia de alumnado con dificultades de aprendizaje entre los niños (Romero, 1999). Estos apuntes, que tendremos la oportunidad de ver o no confirmados con nuestros datos, podrían influir en que las comparaciones efectuadas estuvieran realizadas con grupos poco equivalentes en razón de la mayor presencia de uno de los sexos, que sesgara hacia arriba o hacia abajo las puntuaciones obtenidas. Hemos realizado diversos contrastes estadísticos con la prueba *chi cuadrado* para evaluar esta equivalencia. Primero la efectuamos para el GE y, posteriormente, para el GC1.

Tabla 4. Prueba de chi-cuadrado para el análisis de la equivalencia en el Grupo Experimental. Frecuencias

	sexo (GE)		Residual
	N observado	N esperado	
niñas	23	30,0	-7,0
niños	37	30,0	7,0
Total	60		

Estadísticos de contraste	
	Sexo
Chi-cuadrado ^a	3,267
Gl	1
Sig. asintót.	,071

a. 0 casillas (,0%) tienen frecuencias esperadas menores que 5. La frecuencia de casilla esperada mínima es de 30,0.

Tabla 5. Prueba de chi-cuadrado para el análisis de la equivalencia en el Grupo Control. Frecuencias

	sexo (GC1)		Residual
	N observado	N esperado	
niñas	29	36,5	-7,5
niños	44	36,5	7,5
Total	73		

Estadísticos de contraste	
	Sexo
Chi-cuadrado ^a	3,082
GI	1
Sig. asintót.	,079
a. 0 casillas (,0%) tienen frecuencias esperadas menores que 5. La frecuencia de casilla esperada mínima es de 36,5.	

Como puede comprobarse, tanto en el GE como en el GC, hay una mayor presencia de niños con dificultades de aprendizaje en su lectura, a tenor de la información otorgada por los profesores-colaboradores, en comparación con las niñas que muestran estos mismos problemas. Concretamente, los números totales son de 81 niños y 52 niñas. Este mayor número de niños en nuestra muestra, que como recordábamos anteriormente ha sido documentado en reiteradas ocasiones cuando se ha abordado el estudio de los problemas de aprendizaje, no llega a traducirse en diferencias estadísticamente significativas y mantiene una proporción muy similar entre los grupos experimental y control con respecto a la distribución muestral por sexos; aspecto este último por el que el posible sesgo en los resultados debido a este factor parece estar controlado.

Dado que algunos de los contrastes los efectuaremos diferenciando las etapas de educación primaria y secundaria, hemos optado por realizar esta prueba de equivalencia de forma separada para ambas submuestras. Por lo que respecta a educación primaria, las diferencias observadas anteriormente en el análisis global se mantienen para el GE, sin llegar a ser significativas; en cambio, para el GC estas diferencias pasan a ser mínimas. Aunque, en cualquier caso, la comparación entre ambos grupos parece legítima, el mayor número de niñas en el GC de primaria, de confirmarse mejores resultados en las pruebas de comprensión y en las actividades del dispositivo de evaluación dinámica, podría introducir cierto sesgo a favor de este grupo.

Tabla 6. Prueba de chi-cuadrado para el análisis de la equivalencia en primaria para los grupos experimental y control. Frecuencias

	sexo (GE y GC)					
	N observado		N esperado		Residual	
	GE	GC	GE	GC	GE	GC
niñas	12	17	16,0	18,0	-4,0	-1,0
niños	20	19	16,0	18,0	4,0	1,0
Total	32	36				

Estadísticos de contraste		
	Sexo	
	GE	GC
Chi-cuadrado ^a	2,000	0,111
GI	1	1
Sig. asintót.	0,157	0,739
a. 0 casillas (,0%) tienen frecuencias esperadas menores que 5. La frecuencia de casilla esperada mínima es de 16,0. para el GE y de 18 para el GC.		

Al igual que ocurría para el grupo de alumnos y alumnas de primaria, en secundaria, el GE mantiene una proporción similar a la observada en el análisis global; en cambio, en el GC la presencia de niños es superior al doble de las niñas, siendo en este caso

significativas las diferencias observadas. En este sentido, al igual que anotamos anteriormente, solo que en sentido inverso, debemos estar especialmente atentos a los resultados obtenidos por el grupo de chicas y por el de chicos, por si se estuviera sesgando la comparación entre los grupos experimental y control debido a una presencia significativamente mayor de niños en el GC.

Tabla 7. Prueba de chi-cuadrado para el análisis de la equivalencia en secundaria para los grupos experimental y control. Frecuencias

sexo (GE y GC)						
	N observado		N esperado		Residual	
	GE	GC	GE	GC	GE	GC
niñas	11	12	14,0	18,5	-3,0	-6,5
niños	17	25	14,0	18,5	3,0	6,5
Total	28	37				

Estadísticos de contraste

	Sexo	
	GE	GC
Chi-cuadrado ^a	1,286	4,568
GI	1	1
Sig. asintót.	0,257	0,033

a. 0 casillas (,0%) tienen frecuencias esperadas menores que 5. La frecuencia de casilla esperada mínima es de 14,0. para el GE y de 18,5 para el GC.

Por último, realizaremos este análisis en función del número de niñas y de niños que conforman los grupos experimental y control. En este caso, la comparación debe observarse horizontalmente. Un número de niñas o de niños significativamente superior en alguno de estos grupos con respecto al otro, podría introducir un efecto no deseado en los resultados. Los datos reflejan que, entre las niñas, el número de las que forman parte del GC es superior al que pertenecen al GE, y que, entre los niños, igualmente el número de los que conforman el GC es superior a los que forman el GE. Estos análisis, sin ser significativos, si parecen alertar acerca de un cierto sesgo a favor del grupo control en el caso de confirmarse que las niñas obtengan mejores resultados en las pruebas, aunque esto podría quedar compensado por la presencia, también mayoritaria de niños en este mismo GC. Como hemos ya expuesto en las anteriores comparaciones, deberemos observar con especial atención los resultados que obtiene el grupo de chicas y de chicos para determinar si éstos pueden haber influido diferencialmente en los grupos de tratamiento y control.

Tabla 8. Prueba de chi-cuadrado para el análisis de la equivalencia en los grupos experimental y control. Frecuencias

sexo (GE y GC)						
	N observado		N esperado		Residual	
	GE	GC	GE	GC	GE	GC
niñas	23	29	26,0	26,0	-3,0	3,0
niños	37	44	40,5	40,5	-3,5	3,5
Total	60	73				

Estadísticos de contraste				
	Nivel escolar			
	2º ciclo (p)	3º ciclo (p)	1º ciclo (s)	2º ciclo (s)
Chi-cuadrado ^a	2,882	0,176	4,900	1,000
Gl	1	1	1	1
Sig. asintót.	0,090	0,674	0,027	0,317

a. 0 casillas (,0%) tienen frecuencias esperadas menores que 5. La frecuencia de casilla esperada mínima es de 8,5 para el contraste GE/GC 1º ciclo (p); 25,5 para el 2º ciclo (p); 20 para el 1º ciclo (s) y 12,5 para el 2º ciclo (s).

Tan solo resaltar las diferencias significativas encontradas para el contraste del primer ciclo de educación secundaria, en el que la presencia de alumnos del GC1 representa más del doble que la del GE. Sin embargo, en nuestro estudio, un mayor peso que el curso escolar en las posibles diferencias entre los grupos experimental y control, puede tenerlo el hecho de que estos grupos estén en definitiva compuestos por alumnos y alumnas con diferentes tipos y grados de necesidades educativas especiales ligadas a los procesos de lectura y comprensión. En este sentido, nos ha parecido relevante establecer una comparación entre la cantidad de niños y niñas con discapacidad (dis) que están presentes en los grupos experimental y control. Como decíamos anteriormente, la totalidad del alumnado de la muestra presenta dificultades en su aprendizaje relacionadas con la lectura y, con relación a esto, no existirían diferencias significativas en cuanto al número de alumnos que conforman la muestra; en cambio, entre estos alumnos, hay algunos que presentan discapacidad, y un número significativamente superior de ellos en uno de los grupos del estudio, podría llegar a introducir sesgos importantes habida cuenta de que, supuestamente, su rendimiento en las pruebas sería menor. Hemos realizado igualmente estas comparaciones para las variables *dificultades de aprendizaje* (da) (filtrando de la muestra aquellos alumnos y alumnas que solo fueron catalogados dentro de esta categoría, es decir, en ausencia de discapacidad o deprivación), *deprivación socio-cultural* (dep) y *recibe apoyo a la integración* (r.ap.).

Tabla 10. Prueba de chi-cuadrado para el análisis de la equivalencia en nivel escolar en los grupos experimental y control. Frecuencias

	N observado				N esperado				Residual			
	da	dis	dep	r.ap.	da	dis	dep	r.ap.	da	dis	dep	r.ap.
GE	31	13	24	39	39,0	14,5	18,5	30,5	-8,0	-1,5	5,5	8,5
GC1	47	16	13	22	39,0	14,5	18,5	30,5	8,0	1,5	-5,5	-8,5
Total	78	29	37	61								

Estadísticos de contraste				
	Nivel escolar			
	Dificultades de aprendizaje	Discapacidad	Deprivación socio-cultural	Recibe apoyo
Chi-cuadrado ^a	3,282	0,310	3,270	4,738
Gl	1	1	1	1
Sig. asintót.	0,070	0,577	0,071	0,030

a. 0 casillas (,0%) tienen frecuencias esperadas menores que 5. La frecuencia de casilla esperada mínima es de 39,0 para el contraste GE/GC da; 14,5 para dis.; 18,5 para dep. y 30,5 para r. ap.

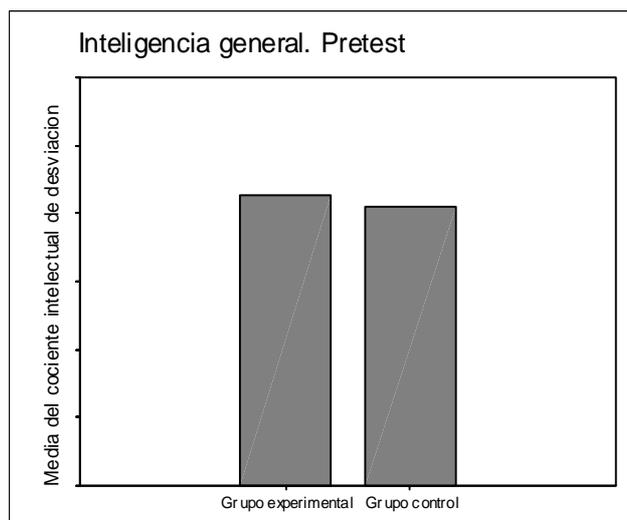
Puede observarse que no existen diferencias significativas en cuanto a la presencia de alumnado con discapacidad en los grupos experimental y de control; sin embargo, las

diferencias son ya notables en cuanto a la cantidad de alumnos y alumnas con DA que forman parte del GC1 con respecto al GE, y, en sentido inverso, con relación al alumnado en situación de desventaja socio-cultural. Estas diferencias llegan a ser estadísticamente significativas para la variable *Recibe apoyo*, configurándose el GE con un número de niños y niñas que reciben apoyo a la integración significativamente superior al del GC1. El hecho constatado de que el alumnado que recibe apoyo a la integración presenta un nivel de dificultades de aprendizaje superior al que recibe otros tipos de apoyo o refuerzo (apoyo a ciclo, refuerzo educativo o agrupamientos flexibles), podría suponer de forma indirecta una falta de equivalencia entre los grupos. Por otra parte, una vez realizados estos mismos contrastes por etapas, los resultados de educación primaria muestran igualmente diferencias significativas: 19 alumnos del GE recibían apoyo, mientras que del GC1, solo 8; la chi-cuadrado resultante para este contraste fue de 4,481 con una significación asociada de 0,034. En cambio, en secundaria, los resultados muestran la práctica ausencia de diferencias: 19 alumnos del GE reciben apoyo a la integración y 18 por parte del GC1.

Equivalencia de nivel intelectual

Aunque no es del todo habitual en los estudios relacionados con la evaluación de la comprensión que se tome en consideración el rendimiento en pruebas de inteligencia general estandarizadas, sí parece bastante extendida la idea, por otra parte, de que el CI de los sujetos puede llegar a ser una variable con cierto peso explicativo en el rendimiento posterior en pruebas con contenido escolar, entre ellas las pruebas de lectura que no se encuentren específicamente ligadas a los procesos de asociación de grafemas y fonemas. El control del nivel intelectual de los grupos experimental y control se ha efectuado mediante la valoración realizada con una prueba de factor «G» estandarizada, *Cattell 2-Forma A*. Los resultados han sido tipificados en puntos CI con media 100 y desviación típica 16. Lógicamente, no debíamos hallar diferencias significativas de partida entre los grupos para considerar controlada esta variable.

Figura 3. Comparación de puntuaciones pretest en CI para los grupos experimental y control



Los resultados de la aplicación de la prueba estadística t de Student para muestras independientes nos muestran la ausencia de diferencias estadísticamente significativas entre las puntuaciones de partida obtenidas en Cattell por los sujetos de ambos grupos. El gráfico muestra la similitud en las puntuaciones medias. La hipótesis nula, expresada

en los términos de que los grupos experimental y control pertenecen a poblaciones distintas, queda rechazada y, por tanto, no hay motivos para pensar que el desigual nivel intelectual de partida de ambos grupos sea relevante en la explicación de cambios posteriores.

Tabla 11. Cociente intelectual de desviación. Puntuación obtenida en la fase pretest en la prueba de inteligencia general de Cattell 2 - Forma A

	N	Media	DT	Error típico	Mínimo	Máximo
GE	60	82,57	14,00	1,81	52	109
GC1	45	80,91	18,05	2,69	44	114
Total	105	81,86	15,80	1,54	44	114

Prueba t de muestras independientes

	Prueba Levene		Prueba t para la igualdad de medias		
	F	Sig.	t	Gl	Sig. (bilateral)
Se han asumido varianzas iguales	3,787	,054	,529	103	,598
No se han asumido varianzas iguales			,511	80,45	,611

El nivel intelectual medio de ambos grupos es medio-bajo, rondando en ambos casos el límite de una desviación típica por debajo de la media. Con carácter complementario, hemos comparado las puntuaciones de los sujetos con discapacidad que se encontraban en los grupos experimental y control. Si antes realizábamos esta comparación teniendo en cuenta el número de estos sujetos que integraba cada uno de los grupos, comprobando que no había diferencias significativas en cuanto a su presencia en ambos, ahora lo hacemos teniendo en cuenta sus puntuaciones en la prueba de inteligencia. Unas puntuaciones especialmente bajas por parte de uno de los subgrupos de sujetos con discapacidad en uno de los grupos, podría introducir un sesgo difícil de controlar, que, por otra parte, haría más compleja la interpretación de nuestros resultados. Igualmente, hemos realizado estas comparaciones para los subgrupos de sujetos con dificultades de aprendizaje, desventaja socio-cultural y para aquellos que reciben apoyo a la integración. En cuanto al alumnado con discapacidad, los resultados muestran que no existen diferencias estadísticamente significativas entre ambos grupos con respecto a las puntuaciones obtenidas en Cattell por estos alumnos. El valor medio para el GE ha sido de 76,31 (DT = 10,76), mientras que para el GC1 ha sido de 68,25 (DT = 14,30); pese a esta diferencia en las medias, tanto el valor de la prueba de homogeneidad de varianzas ($F = 0,528$; sig. 0,474), como el valor de la t para 27 gl. ($t = 1,680$; sig. 0,105) indican la ausencia de significatividad en estas diferencias y, por tanto, garantías suficientes en comparaciones posteriores. Por lo que respecta al subgrupo de alumnos y alumnas con dificultades de aprendizaje, los resultados muestran igualmente ausencia de diferencias significativas. La media de los 31 sujetos del GE que realizaron la prueba fue de 83,06 (DT = 15,03), mientras que la de los 24 del GC1 fue de 87,71 (DT = 16,36). La prueba t para muestras independientes obtuvo un valor de $-1,093$, con una significación asociada de 0,279; la prueba de homogeneidad de varianzas dio un valor F

de 0,454 con una significación de 0,504. En cuanto a la desventaja social, los resultados muestran igualmente que no se dan diferencias significativas. El valor medio de los 24 sujetos del GE fue de 82,63 (DT = 14,14) y el de los 8 sujetos del GC1 de 80,75 (DT = 17,36). La prueba t para muestras independientes obtuvo un valor de 0,307, con una significación asociada de 0,761; la prueba de homogeneidad de varianzas dio como resultado un valor F de 0,378, con una significación de 0,543. En definitiva, para este subgrupo tampoco podemos concluir que existan diferencias de partida que puedan introducir un sesgo en los resultados posteriores. Por último, en cuanto al alumnado que recibe atención educativa especial en las aulas de apoyo a la integración, nuestras comparaciones arrojaron unos resultados que, al igual que ocurría cuando comparamos el número de sujetos presentes en cada grupo que recibía este tipo de apoyo, requerirán de especial atención por nuestra parte en el posterior análisis de los resultados. La puntuación promedio de los 39 sujetos del GE que recibieron este tipo de apoyo fue de 79,82 (DT = 13,37), mientras que en el GC1 fue de 70,70 (DT = 14,61). Esta diferencia llega a ser significativa al 0,05: $t = 2,404$; sig. 0,019 para 57 gl. El valor de la prueba de homogeneidad de varianzas fue: $F = 0,183$, con una sig. asociada de 0,670. Los resultados hallados para este subgrupo parecen a primera vista contradecir nuestra suposición de que una mayor cantidad de alumnos que recibían apoyo a la integración en el GE podría introducir un sesgo a favor del GC1, ya que, en condiciones normales, estos sujetos requieren una atención más específica porque tienen más dificultades de aprendizaje; en cambio, esta suposición sigue teniendo pleno sentido por cuanto las comparaciones post-tratamiento tendrán lugar sobre todo entre los grupos experimental y control al completo. En estas comparaciones formarán un solo grupo todos los alumnos y las alumnas con dificultades de aprendizaje de la lectura, es decir, incluyendo todos los subgrupos que hemos estado viendo, independientemente de que su grado de dificultad observe notables diferencias; es decir, que, como hemos podido comprobar, a tenor de las diferencias existentes entre las puntuaciones obtenidas por el alumnado que asiste a las aulas de apoyo y aquel que no recibe esta atención específica, una mayor presencia de estos alumnos en uno de los grupos podría seguir suponiendo un handicap en su comparación con el otro. Lógicamente, a tenor de los resultados que hemos obtenido en estos primeros contrastes, efectuaremos igualmente comparaciones por subgrupos, teniendo en cuenta factores relativos al grado de dificultad existente en función de los diversos criterios a los que hacíamos alusión al comienzo de este capítulo. Estas comparaciones matizarán, al tiempo que enriquecerán, las conclusiones que podamos obtener sobre la aplicación del dispositivo.

Una última consideración con referencia a la muestra principal de alumnos con dificultades de aprendizaje. La lógica y la naturaleza de los datos manejados nos hacen constatar la existencia de una dispersión de estos datos con relación al valor promedio. Dado que hemos visto cómo las puntuaciones obtenidas por algunos de los subgrupos que conforman la muestra, podían desviarse notablemente de la media, hemos optado por realizar un contraste entre los grupos experimental y control para evaluar el ajuste de las distribuciones de las puntuaciones CI. Para ello hemos utilizado la prueba no paramétrica de la Mediana, a través de la cual puede obtenerse un estadístico chi-cuadrado para comparar la distribución de las puntuaciones CI en los dos grupos. El valor obtenido ha sido de 1,680, con una significación asociada de 0,195. Estos valores confirman que, a nivel de conjunto, los grupos son equivalentes con relación al nivel intelectual, no solo teniendo en cuenta los valores promedio, sino también en cuanto a la distribución de las puntuaciones.

Equivalencia con relación al rendimiento en lectura

Una investigación interesada en la aplicación y valoración de un procedimiento de evaluación de la comprensión de textos, debe contemplar de algún modo la posibilidad de que los grupos experimentales que le sirven al propósito de establecer los posibles cambios, no sean totalmente equivalentes en cuanto al nivel que presentan en tareas de lectura y comprensión. Se hace necesario, por tanto, establecer algún tipo de control para garantizar en lo posible que los resultados obtenidos tras la aplicación del dispositivo de evaluación dinámica no se deban a la existencia de desfases previos en los niveles de pericia en la actividad de comprender textos. El nivel de rendimiento previo en lectura ha sido controlado mediante la utilización de las pruebas ECO1, en educación primaria, y ECOS en secundaria. Obviamente, para considerar controlada esta variable no debíamos encontrar diferencias estadísticamente significativas tras el análisis de las puntuaciones iniciales, obtenidas en las pruebas citadas por el alumnado participante y perteneciente a los grupos experimental y de control. Para el subgrupo de educación primaria, los resultados aparecen en la siguiente tabla.

Tabla 12. Prueba ECO 1 de evaluación de la comprensión de textos en primaria. Puntuación obtenida en la fase pretest

	N	Media	DT	Error típico	Mínimo	Máximo
GE	32	84,18	8,7468	1,5462	66,73	109,35
GC1	36	86,16	8,9071	1,4845	68,97	104,86
Total	68	85,23	8,8227	1,0699	66,73	109,35

Prueba t de muestras independientes

	Prueba Levene		Prueba t para la igualdad de medias		
	F	Sig.	t	Gl	Sig. (bilateral)
Se han asumido varianzas iguales	0,254	0,616	-0,925	66	0,358
No se han asumido varianzas iguales			-0,926	65,32	0,358

Los resultados mostrados avalan la equivalencia de ambos grupos en cuanto a sus puntuaciones de partida en la prueba de comprensión realizada. Esta equivalencia queda reforzada por el valor obtenido en la prueba de homogeneidad de varianzas, por lo que la comparación entre ambos grupos es lícita a tenor de estos análisis y, por consiguiente, las puntuaciones post-tratamiento en esta misma prueba no deberían ser atribuidas a condiciones desiguales con relación al nivel de comprensión.

Al igual que hemos hecho en las anteriores comparaciones, nos interesa saber cómo se han comportado los subgrupos que hemos estado considerando para estos contrastes. Los resultados de los análisis muestran la ausencia de diferencias estadísticamente significativas tanto para el subgrupo de alumnos con discapacidad ($t = 0,756$; sig. 0,469), para el conformado por el alumnado que presenta sólo dificultades de aprendizaje de la lectura ($t = -0,581$; sig. 0,566), con relación al compuesto por

alumnado con deprivación socio-cultural ($t = 0,105$; sig. 0,917), como para los que reciben apoyo a la integración ($t = 1,190$; sig. 0,247). En todos estos contrastes se han obtenido niveles altos de homocedasticidad, por lo que todo parece indicar la equivalencia con relación al rendimiento previo en lectura para el alumnado de primaria, tanto a nivel global como considerando los diversos subgrupos que componen nuestra muestra.

Con respecto al grupo de alumnos y alumnas de secundaria, los resultados que arrojan los análisis globales son estos:

Tabla 13. Prueba ECOS de evaluación de la comprensión de textos en secundaria. Puntuación obtenida en la fase pretest

	N	Media	DT	Error típico	Mínimo	Máximo
GE	28	88,25	10,2412	1,9354	73,73	111,35
GC1	37	91,17	8,0258	1,3194	75,18	105,56
Total	65	89,91	9,0886	1,1273	73,73	111,35

Prueba t de muestras independientes

	Prueba Levene		Prueba t para la igualdad de medias		
	F	Sig.	t	Gl	Sig. (bilateral)
Se han asumido varianzas iguales	1,669	0,201	-1,289	63	0,202
No se han asumido varianzas iguales			-1,247	49,85	0,218

En cuanto a los análisis efectuados para los distintos subgrupos, los resultados muestran que no existen diferencias significativas para el subgrupo de DA ($t = -0,635$; sig. 0,529); así mismo, para el subgrupo con deprivación socio-cultural las diferencias son inapreciables ($t = 0,077$; sig. 0,942); en cambio sí se dan diferencias estadísticamente significativas a favor del GC1 en el subgrupo de alumnos y alumnas que presentan discapacidad en secundaria. Este subgrupo estaba compuesto por 6 alumnos en el GE y 12 en el GC1, aunque el contraste con la prueba chi-cuadrado no reveló diferencias estadísticamente significativas en cuanto a esta mayor presencia de sujetos con discapacidad en el GC1 (chi-c.= 2,000; sig. 0,157 para 1 grado de libertad); la t resultante para las puntuaciones fue de $-2,585$ y la significación asociada de 0,020. Por su parte, el subgrupo de alumnos y alumnas que reciben apoyo a la integración volvió a mostrar diferencias significativas entre GE y GC1 a favor del segundo. En este caso, el número de sujetos que integran cada uno de los grupos es muy similar: 19 forman parte del GE y 18 lo hacen del GC1; la t arroja un valor de $-2,383$ y la significación asociada es de 0,023. En todos los contrastes realizados, los valores de la prueba de homogeneidad de varianzas indican que éstas no difieren significativamente entre los grupos experimental y control.

Equivalencia con relación al ajuste personal-social

El nivel de adaptación o ajuste personal-social es un factor relevante a la hora de explicar el éxito escolar o académico y, por tanto, también el éxito de cualquier programa de intervención en el ámbito escolar. Igualmente, en un estudio como el que presentamos, la valoración de los procesos de ajuste personal y social implicados en la lectura, que manifiestan los sujetos participantes, se torna esencial por cuanto son considerados desde nuestra perspectiva, y también desde la de muchos otros investigadores citados en nuestro estudio, como procesos clave que pueden ayudarnos a determinar o precisar dónde se encuentran las dificultades del sujeto y cómo debemos actuar para tratar de solucionar los problemas. La equivalencia inicial de los grupos experimental y control en cuanto a la variable ajuste personal-social, se ha realizado mediante la escala APSL (Ajuste Personal-Social con relación a la Lectura). La ausencia de diferencias significativas en los procesos evaluados con esta escala, para los dos grupos, nos daría mayores garantías de que el nivel de partida en cuanto a esta variable no introduce sesgos a la hora de valorar e interpretar los posibles cambios posteriores. Los resultados para el grupo completo están expresados en la siguiente tabla.

Tabla 14. Prueba APSL de ajuste personal-social con relación a la lectura. Puntuación obtenida en la fase pretest

	N	Media	DT	Error típico	Mínimo	Máximo
GE	57	97,98	13,9133	1,8429	71,21	122,13
GC1	60	95,06	15,9073	2,0536	54,23	124,96
Total	117	96,39	14,96	1,3837	54,23	124,96

Prueba t de muestras independientes

	Prueba Levene		Prueba t para la igualdad de medias		
	F	Sig.	T	Gl	Sig. (bilateral)
Se han asumido varianzas iguales	0,157	0,693	0,982	115	0,328
No se han asumido varianzas iguales			0,986	114,2	0,326

Como puede comprobarse, no existen diferencias que puedan considerarse relevantes desde un punto de vista estadístico. El valor de la prueba de homogeneidad de varianzas también es indicativo de que ambos grupos son equivalentes de partida en cuanto al nivel de ajuste personal-social manifestado a través de la prueba APSL. Hemos realizado este mismo contraste para primaria y para secundaria. La prueba, como más tarde veremos, aunque ha sido considerada totalmente equivalente para ambas etapas, contenía algunas variaciones en la redacción de algunos ítems con el objetivo de facilitar la comprensión del alumnado sobre el contenido esencial de las cuestiones. Los contrastes para primaria muestran ausencia de diferencias significativas entre el GE y el GC1. El valor medio de los 30 alumnos del GE fue de 98,6987 (DT = 13,4798) y el de

los 23 del GC1 de 96,4252 (DT = 19,5202); el valor de la t fue de 0,501 para 51 grados de libertad, y la significación asociada fue de 0,618. Para secundaria, el contraste arroja igualmente ausencia de significatividad en las diferencias medias. El valor promedio de los 27 alumnos que formaban parte del GE fue de 96,7760 (DT = 14,5688), mientras que para el GC1 el valor fue de 94,2243 (DT = 13,4142); la t observada fue de 0,725 y la significación para 62 grados de libertad fue de 0,471. En ambos contrastes, los niveles de homocedasticidad mostraron equivalencia entre las varianzas.

En cuanto a los distintos subgrupos, todos los contrastes realizados muestran una similitud notable entre las puntuaciones y, por tanto, ausencia de diferencias significativas. Por ejemplo, para el subgrupo DA el valor t ha sido de 0,075 y la significación asociada, para 63 grados de libertad, de 0,941. Para el subgrupo discapacidad la t observada ha sido de 0,086 y el valor de significación de 0,932 para 27 gl. Igualmente, aunque con menor igualdad, se ha mostrado el subgrupo de alumnos y alumnas con deprivación; la t ha tenido un valor de 1,548 y la significación ha sido de 0,132 para 32 gl. Por último, el subgrupo que recibe apoyo a la integración ha mostrado esta vez una total equivalencia en sus puntuaciones, con un valor t de 0,161 y una significación para 58 gl. de 0,873. En todos estos contrastes los valores de la prueba de homogeneidad de varianzas indica ausencia de diferencias significativas entre las varianzas para los grupos de la investigación.

Conclusiones sobre la equivalencia entre los grupos

En definitiva, del análisis de la totalidad de contrastes realizados podemos concluir que, a nivel general, es decir, tomando en consideración la muestra en su conjunto, la equivalencia entre el grupo experimental y el grupo control es notable y, por tanto, la comparación entre ambos es lícita, ofreciendo garantías suficientes para establecer, tras la fase de tratamiento, las posibles diferencias entre los sujetos participantes de uno u otro grupo, con relación a la aplicación experimental del dispositivo de ED. Los análisis efectuados para los grupos experimental y control no muestran diferencias significativas en cuanto a la distribución de efectivos por sexo y edad. Con relación al nivel escolar, todos los contrastes a excepción del efectuado en el primer ciclo de secundaria muestran la equivalencia entre los grupos, y cuando las comparaciones se efectúan entre los efectivos de los cuatro subgrupos considerados (DA, discapacidad, deprivación y alumnado que recibe apoyo a la integración), tan solo el grupo *recibe apoyo* muestra diferencias significativas entre el GE y el CG1 a favor del primero. Los grupos también muestran equivalencia con relación al nivel intelectual, y al igual que para el elemento anterior, el subgrupo formado por el alumnado que recibía apoyo a la integración ha mostrado un desigual desempeño en función de su pertenencia al GE o al GC, también a favor del GE. Así mismo, en cuanto al rendimiento previo en lectura, tanto para primaria —considerando los cuatro subgrupos— como para secundaria —en este caso, con la salvedad de los subgrupos *discapacidad* y *recibe apoyo* cuando han sido analizados por separado, en los que se han hallado diferencias a favor del GC1—, los grupos no difieren significativamente. Por último, en cuanto al nivel de ajuste personal-social, los resultados muestran un grado de equivalencia importante, tanto a nivel general como para las distintas etapas y los distintos subgrupos.

5.4 La muestra de la aplicación piloto

Aunque posteriormente haremos mención de forma más específica al proceso de la aplicación piloto, cuando nos detengamos en la descripción y el análisis de los instrumentos elaborados y utilizados en nuestro estudio, exponemos ahora los aspectos relativos a la selección y composición de esta submuestra.

Como apuntábamos en páginas anteriores, la necesidad de realizar una aplicación piloto de las pruebas elaboradas para poder operativizar posteriormente las variables relativas al rendimiento en tareas de lectura y al nivel de ajuste personal-social con relación a la lectura, provocó la búsqueda de un grupo de sujetos suficiente y representativo en lo posible (en cuanto a edad, contexto socio-cultural y nivel educativo) del conjunto de alumnos y alumnas que formarían parte de la investigación. Un aspecto que debemos reseñar es que se optó por realizar la aplicación piloto del estudio con grupos clase completos y no específicamente con alumnado que presentara dificultades de aprendizaje. Las pruebas elaboradas, aunque inicialmente pensadas para ser aplicadas a sujetos con especiales dificultades en el aprendizaje de la lectura, una vez construidas pretendían dar cuenta del abanico de capacidades que podían mostrar los sujetos participantes, desde aquellos que presentaban dificultades en su lectura hasta aquellos que mostraran un alto nivel de lectura y comprensión. De esta forma, el análisis de los resultados podría ayudarnos a situar mejor al alumnado con dificultades que formaba parte de los grupos experimental o control dentro de un contexto en el que también existía información sobre la realización de compañeros y compañeras sin especiales dificultades.

La muestra ha estado constituida finalmente por 224 alumnos y alumnas de dos centros educativos de la provincia de Sevilla. 179 sujetos pertenecían al CEIP (centro público de educación infantil y primaria) Manuel Sánchez Alonso, de Arahal (Sevilla) y 45 sujetos pertenecían al IES Dña Leonor de Guzmán, en Alcalá de Guadaíra, también en Sevilla. Por otra parte, la muestra la formaron, como anotábamos anteriormente, alumnos pertenecientes a distintos grupos-clase completos de educación primaria y secundaria. La selección de los grupos se ha realizado en función de su pertenencia al rango de edad y nivel educativo estipulado en el estudio y, obviamente, en función de la mayor o menor facilidad que otorgara el centro educativo y cada uno de los grupos. Al realizar la aplicación con grupos completos, lógicamente, también hay alumnos y alumnas con necesidades educativas especiales y dificultades de aprendizaje ligadas a la lectura. Se ha optado por incluir a todos los sujetos pertenecientes al grupo-clase que se tratara (salvo que el alumno no pudiera materialmente realizar la prueba debido a dificultades graves que le impidieran acceder a la comprensión de la tarea) con el objetivo de lograr una mayor representatividad de la muestra y, como consecuencia, unos resultados que fueran indicativos de la totalidad y la diversidad del alumnado presente habitualmente en los centros. Los dos centros educativos, como puede haberse ya comprobado, han participado en el estudio con alumnado que formaba parte de los grupos experimental y control, y también mediante la colaboración especial de profesores y orientadores adscritos a ellos.

La distribución de esta muestra piloto en función del sexo de los participantes, las pruebas aplicadas y los niveles educativos, se encuentra recogida en la siguiente tabla.

Tabla 15. Composición y distribución de la submuestra para la aplicación piloto de las pruebas ECO1, ECOS y APSL

		ECO1	ECOS	APSL
Sexo	Niñas	88	18	43
	Niños	91	25	55
Nivel educativo	4° EP	22		
	5° EP	83		25
	6° EP	74		30
	2° ESO		22	22
	4° ESO		21	21
Totales		179	43	98

6. Instrumentos

Como hemos venido destacando a lo largo de este trabajo, nuestra propuesta consiste en la aplicación de un procedimiento dinámico de evaluación de procesos implicados en la comprensión del lenguaje escrito, especialmente pensado para ser trabajado con sujetos que presentan especiales dificultades en el aprendizaje y desarrollo de su lectura. A su vez, esta propuesta está fundamentada en las teorías y los trabajos que consideran, de una u otra forma, que tanto el aprendizaje en general como la lectura en particular, son procesos que se construyen esencialmente a través de la interacción mediada y, así mismo, que esta acción educativa puede estructurarse y sistematizarse en el ámbito educativo, llegando a contribuir significativamente a que se modifiquen y optimicen determinadas estructuras o funciones cognitivas y socio-personales implicadas en la construcción de nuevos conocimientos, actitudes y habilidades (Vygotski, 1934/1995; Solè, 1992; Coll y Solè, 1993; Mora, 1998; Jensen, 2000).

En línea con lo anteriormente expuesto, parece obvio que esta naturaleza interactiva y dinámica de la lectura y de los procesos de enseñanza y aprendizaje de la misma, debería reflejarse de forma clara en cualquier procedimiento de evaluación que pretenda acceder al proceso de lectura y a los procesos que intervienen en la comprensión de los textos. De igual forma, como consecuencia, y en coherencia con la tesis planteada, el análisis y la evaluación de los estudios sobre la aplicación dinámica de un procedimiento de evaluación, debería contemplar sistemas dinámicos de valoración y análisis, que pudieran acercarse a captar la complejidad y la riqueza de los procesos puestos en juego durante la experiencia. Esto, en cambio, como veíamos cuando abordábamos la revisión crítica de los estudios realizados sobre evaluación dinámica, no ha constituido la praxis habitual en la investigación. Por el contrario, la mayoría de estudios que han abordado procesos de evaluación interactiva, han sido evaluados mediante procedimientos de corte estático y convencional (Lidz y Elliot, 2000; Grigorenko y Sternberg, 1998; Saldaña, 2001; Sternberg y Grigorenko, 2003), que han primado la ejecución individual, autónoma y, en no pocas ocasiones, descontextualizada de las actividades. Esto ha ocurrido, tanto para los casos en los que el contenido de los estudios ha sido la inteligencia o las funciones cognitivas generales, como cuando el objeto de estudio ha sido la evaluación dinámica de funciones cognitivas específicas o relacionadas estrechamente con la actividad escolar. Probablemente, las razones que fundamentan esta decisión podrían sintetizarse, por una parte, en el énfasis puesto en la cuantificación de las medidas obtenidas, con el objeto de tener la posibilidad de realizar comparaciones estándar con otros resultados obtenidos en estudios relacionados; y, por otra, en la posibilidad, que habitualmente le es otorgada a una investigación cuyo análisis se realiza mediante una evaluación psicométrica, de ser replicada por investigadores independientes. En nuestro caso, podríamos decir que algunas de estas razones también han tenido un peso considerable en la adopción final de un procedimiento de análisis que contemplara de forma relevante la utilización de pruebas psicométricas. Posteriormente, cuando nos adentremos en el capítulo de Discusión, profundizaremos en el análisis de estas razones, así como en las posibles consecuencias que haya podido tener la consideración de estos procedimientos en el desarrollo y los resultados de nuestra investigación.

6.1 Instrumentos para las fases pretest y postest

Como hemos recogido en páginas anteriores, se utilizaron algunas pruebas en la fase pretest del estudio con el objeto de obtener resultados que nos permitieran operativizar y decidir si existían o no diferencias significativas de partida, entre los grupos experimental y control, en variables como la inteligencia, el rendimiento en lectura o el ajuste personal-social. Estas pruebas han sido las que finalmente conformaron la batería criterio de nuestro estudio y fueron administradas antes y después de la aplicación del dispositivo, tanto al grupo experimental como al grupo de control. Hemos utilizado las mismas pruebas como pretest y postest. El intervalo temporal entre una pasación y otra ha sido de unos seis meses; posteriormente a la aplicación experimental, las pruebas se aplicaron nuevamente a ambos grupos. En este sentido, a pesar de que las pruebas realizadas por los sujetos en la fase inicial del estudio, han sido las mismas que las que se han aplicado con posterioridad al tratamiento, el tiempo transcurrido entre una fase y otra se ha estimado suficiente para neutralizar los efectos *positivos* de la práctica, entendida como cierto entrenamiento en la resolución de los test, así como también los *negativos*, entendidos como la posibilidad de que la reiteración en la ejecución de la prueba pueda afectar negativamente al sujeto ocasionándole cansancio, fatiga, desidia o falta de motivación (Fontes y Fontes, 2001). Así mismo, el control experimental de esta variable también se ha apoyado fundamentalmente en la existencia de un grupo de control, que ha recibido de igual forma y en la misma fase temporal, las mismas pruebas que el grupo experimental. Por otra parte, se han aplicado otras pruebas durante el desarrollo de la investigación con el objetivo de aumentar la información y las posibilidades de explorar los efectos de la aplicación del tratamiento.

El test de factor «G» de Cattell

Entre las pruebas seleccionadas en función de los objetivos y las hipótesis formuladas, se encontraba una que otorgaba una medida de la inteligencia general: el *test de factor «G» de Cattell, Escala 2-Forma A* (Cattell y Cattell, 1974/2001). Esta prueba pertenece al conjunto de tests denominados «libres de cultura» y, por tanto, más allá de las limitaciones que han sido consideradas y reflejadas también en este trabajo acerca del término y sus connotaciones para la interpretación de la prueba, su uso nos proporcionaría una medida de la llamada por el propio Cattell *inteligencia fluida*, es decir, aquella que parece estar más relacionada con patrones hereditarios y a su vez menos influida por el trabajo desarrollado desde los ámbitos educativos como la escuela. Esta prueba nos ofrecería una medida de la inteligencia que resultaría útil por dos razones principalmente: (a) establecer la idoneidad y equivalencia de partida de la muestra en cuanto al cumplimiento de las condiciones que se exponían en páginas anteriores, relacionadas con la determinación del tipo de alumnos y alumnas con dificultades de aprendizaje de la lectura que eran susceptibles de participar en los distintos grupos que pueden formarse en la investigación; y (b) establecer un control de la variable inteligencia, medida con este test, entre los grupos experimental y control.

La escala 2 del test de factor «G» de Cattell está pensada y baremada para sujetos de 8 a 14 años, así como para adultos de un nivel cultural medio. Las puntuaciones directas de este test son transformadas con el objetivo de mejorar la interpretación de los resultados obtenidos. De esta forma, las puntuaciones obtenidas por los sujetos se transforman en un *cociente intelectual de desviación*, igualando la media en 100 y la desviación típica en 16. Esta transformación de las puntuaciones directas permite realizar comparaciones

de los resultados para distintas edades, así como con otros tests de inteligencia ampliamente utilizados y que han sido tipificados de forma similar. En nuestro caso, ya hemos hecho mención a las edades comprendidas en la muestra que ha participado en el estudio, así como al nivel socio-cultural que ha caracterizado en líneas generales al conjunto de la población de estudio. Teniendo en cuenta estos dos aspectos, consideramos que esta prueba era la más idónea para valorar el nivel intelectual de los sujetos. Por otra parte, buscábamos una prueba que pudiera ser aplicada de forma colectiva, aspecto que facilitaría significativamente la labor de los profesionales que han colaborado en la pasación de las baterías pretest y postest en los centros, al tiempo que relativamente ágil en cuanto a su aplicación y realización por parte del alumnado.

La escala APSL

La utilización de una prueba que evaluara aspectos relativos a la motivación, las actitudes, las propias expectativas, el autoconcepto, y, en definitiva, todos aquellos procesos de adaptación psico-social que parecen tener una influencia significativa en el desarrollo de los procesos de enseñanza y aprendizaje de la lectura (Mora Merchán y Mora, 2000; Moreno y García, 2000; Navarro y Mora, 2003a), estaba plenamente justificada en nuestro estudio, máxime cuando nosotros mismos hemos otorgado una relevancia crucial en la explicación de las dificultades de aprendizaje y de las posibilidades de optimización, a estos procesos.

Esta utilización de una escala de actitudes relacionadas con la lectura (*La escala de Ajuste Personal-Social con relación a la lectura, APSL*) se ha estimado oportuno por tres razones fundamentales: (a) poder establecer comparaciones entre los resultados de esta prueba y los obtenidos durante la aplicación de la evaluación dinámica de los procesos implicados en la lectura, en la que se tienen en cuenta la motivación, los intereses y las actitudes con relación a la lectura; (b) establecer un control de la variable actitud y motivación, medida con esta prueba, entre los grupos experimental y control; y (c) poder establecer comparaciones entre las fases pretest y postest para los grupos experimental y control en cuanto al ajuste obtenido en los elementos valorados en esta prueba y también con relación al rendimiento en la prueba de comprensión de textos.

La escala APSL (anexo III) se compone de ochenta ítems que se presentan al alumnado en una estructura de frases cortas ante las que éste debe mostrar su acuerdo o desacuerdo. Las frases están relacionadas con la lectura y, especialmente, con el modo en que cada uno se enfrenta a las tareas de lectura y comprensión. Hacen referencia a aspectos relacionados con los propios sentimientos y emociones cuando se está leyendo o se tiene que hacerlo, así como cuando es necesario realizar actividades relacionadas con la lectura. El alumno debe mostrar su acuerdo o desacuerdo con la frase propuesta marcando con una cruz la casilla que se corresponda con su opinión. Aunque no hay un tiempo estipulado para la cumplimentación de la prueba, de forma aproximada, el tiempo empleado por los sujetos que han participado en el estudio ha sido de unos 30 minutos. En las instrucciones de la escala, que eran leídas y comentadas por el evaluador previamente a la ejecución de la prueba, se instaba a los sujetos a leer detenidamente y con mucha atención cada una de las frases para realizar de forma adecuada la actividad y para que pudiera ser válida al tiempo que útil para poder ayudarle con la lectura. Se recogía igualmente en estas instrucciones la importancia de pensar sobre uno mismo a la hora de cumplimentar el ejercicio, siendo sincero en las

respuestas. Por último, en esta hoja de instrucciones se recogía un ejemplo de práctica para que el alumnado pudiera familiarizarse con el tipo de prueba que iba a realizar.

Con la escala APSL, en sus dos versiones: APSL 1 y APSL, para Educación primaria y Educación secundaria respectivamente, pretendemos operativizar la variable socio-emocional en relación con el proceso de lectura y las dificultades de su aprendizaje. La escala APSL 1 ha sido diseñada para su aplicación al alumnado de segundo y tercer ciclo de Educación primaria, mientras que APSL es aplicable al alumnado de Educación Secundaria. Las diferencias entre ambas escalas son mínimas, pero hemos considerado oportuno introducir estas pequeñas diferencias que se han concretado en dos aspectos básicamente: la simplificación del mensaje de algunos de los ítems para la escala de primaria y la utilización de los nombres que las materias reciben habitualmente en una y otra etapa educativa. Estas diferencias no son en absoluto relevantes con relación a la estructura de la prueba ni a las claves de corrección de la misma, por cuanto, en función del interés concreto que en cada momento se ha ido perfilando en el transcurso de la investigación, hemos podido considerar una sola muestra de sujetos (al igual que hemos hecho con la prueba de Cattell) o dos submuestras en función de si cursaban educación primaria o secundaria.

Por otra parte, dado el pretendido carácter objetivable de esta operativización de los aspectos de índole socio-emocional, es preciso aclarar que los ítems de la escala evalúan necesariamente el *conocimiento* o la *percepción* que el alumnado ha construido acerca de sus propias actitudes, motivaciones, intereses, o estrategias de relación con los demás en tareas relacionadas con la lectura. Lógicamente estos conocimientos los ha ido construyendo en función de las experiencias que ha tenido relacionadas con el aprendizaje de la lectura y de cómo ha vivido y afrontado esas experiencias. Por tanto, ha sido necesario tener en cuenta una serie de aspectos para optimizar las posibilidades de establecer un criterio de evaluación adecuado, que nos pudiera servir a su vez para poder valorar posteriormente posibles diferencias en la aplicación posttest del instrumento.

En primer lugar, es necesario hacer mención de la variable *deseabilidad social*. Esta puede manifestarse en las decisiones del alumnado al completar la escala, ya que se trata, a fin de cuentas, de mostrar acuerdo o desacuerdo con frases que hacen relación a los comportamientos relacionados con las actividades de lectura en la escuela o en casa. No es aventurado pensar, por tanto, que algunos alumnos pongan especial énfasis en quedar bien ante los profesores, o ante las personas que vayan a leer el cuestionario, al decidir su opinión. Para evitar en lo posible este hecho, o para minimizar sus consecuencias, hemos utilizado algunos de los recursos que, por otra parte, son habitualmente usados en la construcción de este tipo de instrumentos. Estos recursos han sido:

* Hacer especial hincapié, durante la presentación al alumnado de la actividad a realizar, en la importancia de la *sinceridad de sus respuestas*, y en que solo de esta forma tendrá validez el resultado del ejercicio y podrá servir para ayudarles en su lectura.

* Se ha procurado evitar en lo posible la redacción de ítems que contuvieran ideas muy reiteradas en la escuela acerca de la lectura, y que incitaran por tanto de forma clara a una determinada opción de respuesta.

* Aunque se ha procurado que la redacción de la mayoría de los ítems estuviera formulada en sentido positivo, se han introducido algunos ítems cuya redacción esta formulada en forma negativa. Aunque esta formulación puede llegar a dificultar la comprensión de algunos ítems por parte del alumnado, su conveniencia radica, por una parte, en la mayor fuerza de determinadas ideas que se quieren destacar, cuando estas se formulan en modo negativo (p. e.: *No me gusta leer ni estudiar*) y, por otra, en que puede llegar a provocar en mayor medida la reflexión acerca de los propios sentimientos, emociones o actitudes con relación a la lectura, evitando de esta forma respuestas demasiado rápidas e impulsivas.

* Se han introducido diversos ítems a lo largo de la escala cuya respuesta en un mismo sentido —*acuerdo* o *desacuerdo*— por parte del alumnado anularía en cierta forma su validez a efectos de la puntuación obtenida. Estos ítems, aunque presentan una formulación distinta en su redacción, contienen una idea similar o aproximada. En ocasiones esta idea es la misma, pero expresada de forma distinta. En estos casos, la obtención de menor puntuación se daría si el alumno ha mostrado *acuerdo* en una de ellas y *desacuerdo* en otra. Por otra parte, en otras ocasiones se trata justamente de la idea contraria. La ausencia de puntuación en estos ítems vendría dada por haber mostrado *acuerdo* con un ítem e igualmente *acuerdo* con el ítem que expresa una idea contraria e incompatible. Aunque en un principio habíamos contemplado la posibilidad de que se produjera de forma efectiva una anulación de respuestas en función de una marcación de ítems que se correspondiera con las premisas recogidas más arriba, desechamos posteriormente esta opción por una razón de tipo práctico: un análisis exploratorio de los primeros resultados de la aplicación piloto nos hizo ver que las diferencias finales entre la puntuación obtenida por los sujetos cuando se aplicaba la corrección en función de la anulación de respuestas y cuando ésta no se realizaba, no variaban significativamente con respecto a la puntuación obtenida por los sujetos que manifestaban mayor coherencia. La causa parecía estar en que se producía un cierto efecto de compensación en la puntuación final obtenida, por el cual las diferencias finales entre unos alumnos y otros se mantenían. Aquellos alumnos que dejaban ver en el contenido de sus respuestas una falta de coherencia y de conexión entre las ideas manifestadas, obtenían lógicamente menor puntuación que sus compañeros que sí contestaban guiados por unas ideas más claras y un conocimiento más coherente de sus propios procesos de adaptación social y personal; esto se producía debido a que los primeros, en muchos de los pares de ítems que contestaban de forma inconexa o incoherente, no obtenían puntuación o lo hacían con un solo punto.

Un segundo aspecto que es necesario tener en cuenta a la hora de establecer las claves de corrección del instrumento, tiene que ver con la dimensión teórica del mismo. Tal y como apuntábamos anteriormente, los ítems que conforman la escala tratan de valorar en su conjunto el *conocimiento* y la *percepción* que el alumnado tiene sobre sus propias actitudes, emociones, estrategias de afrontamiento, etc. Esto supone que estamos evaluando lo que el alumnado sabe, piensa y dice en relación con sus propias actitudes, emociones y estrategias; pero hay un componente práctico que, lógicamente, no puede ser valorado con esta escala. Lógicamente, el instrumento ha sido construido para captar en la mayor medida posible estos aspectos mencionados y para que el alumnado, a través de sus decisiones, los deje entrever de la forma más clara posible.

Por otra parte, al valorarse, como hemos expresado, el conocimiento y la percepción del alumnado acerca de los aspectos mencionados, podría considerarse que las respuestas

no tendrían por que conllevar una valoración positiva o negativa. Sin embargo, la constatación en numerosas investigaciones de la relación existente entre el conocimiento sobre los propios procesos de pensamiento, por un lado, y la actividad desarrollada en un aspecto determinado, por otro, nos lleva a establecer una valoración positiva o negativa asociada a cada respuesta dada por el alumnado; valoración basada en los criterios más comúnmente compartidos en relación con el desarrollo y aprendizaje óptimos de los procesos de lectura y comprensión, así como de los procesos de índole socio-emocional que se relacionan con ésta (Mora Merchán y Mora, 2000; Mora, 1998; Moreno y García, 2000; Mateos, 2001; Miras, 2001; Repetto, Téllez y Beltrán, 2002). De esta forma, cada ítem lleva asociada una valoración positiva o negativa y, en consecuencia, una puntuación en función de si la respuesta del alumnado muestra acuerdo o desacuerdo. La puntuación directa máxima que puede obtenerse en la prueba es de 80 puntos. Hay 36 ítems de respuesta en los que marcar *acuerdo* vale 1 punto y 44 ítems en los que marcar *desacuerdo* vale 1 punto. En cualquier caso, hay que tener en cuenta que uno de los aspectos relevantes sería evaluar si existe ajuste/correlación entre la puntuación obtenida en APSL y la puntuación obtenida en la prueba de comprensión de textos. Por tanto, lo importante no sería sólo obtener una puntuación alta en la escala, sino que pueda establecerse una relación entre esta puntuación alta y los resultados de la prueba de comprensión. Si lo que se incrementa entre las fases pretest y postest es precisamente la relación entre las puntuaciones obtenidas en ambas pruebas, esto sería indicativo de mejora y de mayor ajuste.

Por otra parte, la escala APSL contempla una serie de dimensiones relativas a los distintos procesos en que hemos estructurado el ajuste personal-social. Estas dimensiones, como puede comprobarse, han sido consideradas dentro de un marco teórico que contempla la metacognición no sólo como un aspecto perteneciente exclusivamente a la esfera cognitiva y que se relaciona con acciones y aprendizajes de tipo cognitivo o escolar; además de estas consideraciones, la metacognición, en especial la reflexión metacognitiva, pero también los procesos de control y regulación cognitivos, podrían estar en la base de procesos de adaptación social y personal (Mora, 1997, 1998; Mora Merchán y Mora, 2000; Mateos, 2001), permitiendo y facilitando la extensión hacia los demás de un mayor y mejor conocimiento de uno mismo y de las estrategias de afrontamiento de situaciones sociales y de relación. En este sentido el conocimiento de uno mismo y de los propios procesos —cognitivos y de ajuste personal-social— incidirían positivamente en la mejora de la adaptación social y personal de los sujetos. Así mismo, hemos contemplado en la escala una dimensión relativa al metaconocimiento estratégico con relación a la lectura. La razón para esto, cuando, por otra parte, existe ya una prueba específica para evaluar la comprensión de textos en la que se incluyen cuestiones que hacen referencia a los metaconocimientos sobre la comprensión, ha sido doble; por una parte, establecer una valoración alternativa y complementaria de estos procesos que pudiera servir para cotejar y validar los resultados y, por otra, otorgar a la escala construida una estructura y un contenido que contemplara con relación a la lectura todos los aspectos relativos a la comprensión metacognitiva, es decir, todos los elementos relacionados con los procesos de ajuste personal-social, completándola a su vez con una valoración complementaria de los metaconocimientos sobre la comprensión de textos, posibilitando que la puntuación final reflejara de una forma más compleja y completa la situación social y personal del sujeto, así como su conocimiento metacognitivo. En el cuadro siguiente se detalla la estructura de la escala APSL con las distintas dimensiones contempladas, y en el volumen de anexos (Anexo III-A) las claves que permiten corregir la escala.

Cuadro 24. Escalas APSL 1 y APSL (escala de ajuste personal-social con relación a la lectura)

DIMENSIONES DE LA ESCALA

COMPRESION METACOGNITIVA

Conocimiento de uno mismo y de los demás

Conocimiento de las propias limitaciones cognitivas actuales, potencialidades y posibilidades de aprendizaje, dificultades existentes y vías de solución

ítems: 1, 2, 17, 22, 30, 39, 41, 53, 72, 76

Conocimiento de las propias expectativas, autoconcepto académico y estilo atribucional

ítems: 4, 6, 15, 26, 42, 46, 52, 59, 66

Conocimiento de la propia actitud ante la lectura: emociones, sentimiento de rechazo vs satisfacción y gusto por la lectura, tolerancia a la frustración, persistencia ante la dificultad, autonomía vs dependencia

ítems: 7, 8, 11, 14, 18, 24, 28, 34, 36, 40, 50, 51, 55, 69, 71, 73, 78, 80

Conocimiento de los propios intereses y motivaciones

ítems: 10, 13, 19, 32, 38, 45, 57, 68, 70, 75

Percepción del apoyo que ofrece el contexto escolar y familiar, y de las condiciones en las que se desenvuelve

ítems: 3, 21, 27, 35, 37, 44, 48, 56, 58, 62, 65, 67, 74, 77

Percepción de la relevancia del tema

ítems: 12, 47, 54, 60, 79

Metaconocimiento estratégico

Conocimiento de las propias estrategias de relación con los demás en tareas de lectura

ítems: 31, 63, 64

Conocimiento de las propias estrategias de control y regulación consciente del proceso de lectura (planificación, supervisión y auto evaluación)

ítems: 5, 9, 16, 20, 23, 25, 29, 33, 49

Reconocimiento de la capacidad de generalización de estos conocimientos o estrategias

ítems: 43, 61

Cada dimensión de las contempladas en la escala informaría del *ajuste* del alumno con relación a los contenidos valorados. Por ejemplo, una puntuación baja en la dimensión 1, sobre todo cuando no hay incoherencia en la respuesta con respecto a otro ítem de la escala, informaría del conocimiento del sujeto sobre las propias limitaciones cognitivas actuales, las potencialidades y posibilidades de aprendizaje, las dificultades existentes o vías de solución posibles. Habría que comprobar y valorar entonces la correlación entre esta valoración del alumno y sus dificultades reales, evaluadas con la prueba de comprensión y con la valoración del profesorado. En este sentido, podían establecerse varios criterios para evaluar el ajuste personal-social, o algunas de las dimensiones contempladas en la escala, poniendo en relación la puntuación obtenida con los resultados obtenidos en la prueba de comprensión de textos. Estos criterios guiarán parte del análisis de los resultados que efectuaremos posteriormente.

* En primer lugar, se podía establecer una relación entre la puntuación global de la escala y la puntuación obtenida en las dimensiones 8 y 9 de la misma (metaconocimiento estratégico). Se esperaba que las puntuaciones globales más altas tuvieran su correspondencia igualmente en la obtención de mayores puntuaciones en las dos dimensiones citadas. En este sentido, parecía lógico pensar que el alumnado con dificultades de aprendizaje de la lectura obtendría una puntuación global más baja y un mayor desajuste social-personal; así mismo, estas puntuaciones globales más bajas se relacionarían para este grupo con puntuaciones significativamente más bajas en las dimensiones relativas al metaconocimiento estratégico.

* Se podía establecer igualmente, como hemos ya mencionado, una relación entre la puntuación global obtenida en la escala y la puntuación global de la prueba de comprensión.

* Podía evaluarse, así mismo, el ajuste entre la puntuación obtenida en las dimensiones 8 y 9 (metaconocimiento estratégico) de la escala APSL y la puntuación obtenida en la prueba de comprensión.

* También se podía establecer la relación anterior entre la puntuación obtenida en la dimensión de la prueba de comprensión que hace referencia a procesos metacognitivos, y la puntuación obtenida en las dimensiones 8 y 9 de la escala APSL.

Validez y fiabilidad de la prueba APSL

La aplicación piloto de APSL

Dado que la prueba APSL ha sido elaborada para esta investigación, y que con ella pretendía ofrecerse una valoración de los procesos de ajuste personal-social de los sujetos participantes con el objeto de poder establecer un control inicial entre los grupos experimental y control con respecto a la variable socio-emocional, así como diversas comparaciones pretest-postest que nos otorgaran información sobre los posibles cambios en el ajuste personal-social del alumnado; se hacía necesaria una aplicación piloto de la prueba con el objetivo de observar, previamente a su aplicación experimental con los sujetos de la muestra, diversos aspectos relativos a su validez y fiabilidad, así como al tiempo de ejecución estimado, la dificultad o facilidad de los ítems propuestos, la inteligibilidad de las instrucciones planteadas o el interés, la actitud y la motivación con que el alumnado cumplimentaba la prueba.

Esta aplicación piloto, como apuntábamos en páginas anteriores, se ha llevado a cabo en dos centros educativos, el CEIP Manuel Sánchez Alonso, de Arahal y el IES Dña Leonor de Guzmán en Alcalá de Guadaíra, ambos en la provincia de Sevilla. Los participantes han sido dos grupos-clase completos de primaria (55 sujetos), pertenecientes a 5º y 6º curso, y otros dos grupos de secundaria (43 sujetos), de 2º y 4º. Es importante destacar que los grupos-clase han sido seleccionados de centros ya participantes en el estudio y en función de las posibilidades del investigador y la disponibilidad que han mostrado los tutores de los cursos. Dentro de estos grupos, lógicamente, también hay alumnos y alumnas que presentan dificultades de aprendizaje, algunos de los cuales también han participado de nuestro estudio. Hemos optado por mantener a estos alumnos con su grupo a la hora de realizar esta prueba, ya que nos interesaba una muestra que, aunque no muy numerosa, fuera en algo representativa de la diversidad en la población escolar.

En la siguiente tabla se muestra una síntesis de la información estadística básica con relación a la aplicación piloto de la escala APSL.

Tabla 16. Indicadores estadísticos de la aplicación piloto de la escala APSL

Indicadores	Resultados de la aplicación piloto de APSL
Muestra	98
Cursos	5º y 6º EP; 2º y 4º ESO
Edades	10 a 16 años
Ítems de la prueba APSL	80
Índice de Dificultad medio	0,3063
Índice medio de Homogeneidad	0,316
Puntuación media global	104,4369
Error típico de la media	1,5392
Desviación típica	15,2371
Mediana	103,7440
Varianza	232,170
Puntuación mínima	61,31
Puntuación máxima	133,45
Rango	72,14
Amplitud intercuartil	22,6332
Asimetría	-0,382
Curtosis	-0,112

Puntuaciones centiles	5	76,8671
	10	85,3546
	25	95,2566
	50	103,7440
	75	117,8898
	90	123,5481
	95	126,4480
Prueba Kolmogorov-Smirnov de normalidad	0,077 (Sig. 0,179)	
Fiabilidad (índice Alfa de Cronbach)	0,8831	
Etapas		
Media educación primaria	107,6277 (DT = 15,2578; N = 55)	
Media Educación secundaria	100,3556 (DT = 14,3701; N = 43)	
Media chicas	109,0405 (DT = 15,4685; N = 43)	
Media chicos	100,8377 (DT = 14,1730; N = 55)	
Nivel educativo		
Media 5° primaria	110,0248 (DT = 18,8952; N = 20)	
Media 6° primaria	105,6301 (DT = 11,3551; N = 30)	
Media 2° secundaria	103,4869 (DT = 11,7838; N = 22)	
Media 4° secundaria	97,0753 (DT = 16,3023; N = 21)	

Puede observarse que el índice de dificultad medio para este grupo de aplicación piloto en la escala APSL no ha sido alto. Este índice puede ser interpretado como el porcentaje promedio de sujetos que obtienen una valoración negativa en los ítems de la escala; es decir, en nuestro caso, el 30,63% de alumnos. Podríamos comprobar mediante este mismo análisis que este índice se incrementa significativamente cuando valoramos los resultados de los grupos de alumnos y alumnas con dificultades de aprendizaje, lo que nos hablaría, por una parte, de que la muestra de aplicación piloto ha podido ser un buen grupo de referencia para los intereses de nuestro estudio, y, por otra, de la validez de constructo de la prueba, ya que se esperaba que los sujetos que presentaban dificultades de aprendizaje obtuvieran puntuaciones más bajas en la escala. En este sentido, hemos realizado un análisis de varianza con el objetivo de contrastar las puntuaciones del alumnado de la muestra de aplicación piloto (AP) con las del conjunto de alumnos y alumnas con dificultades de aprendizaje que han formado parte, bien del GE o bien del GC1. Los resultados muestran diferencias muy significativas a favor del grupo de sujetos que han participado de la aplicación piloto. Este aspecto confirma nuestra hipótesis y refuerza la validez de la escala APSL por cuanto el alumnado con dificultades en el aprendizaje de la lectura obtiene significativamente una menor puntuación en una medida del ajuste personal-social.

	N	Media	DT	AVAR (F)	gl.	Sig.
Muestra de AP	98	104,4369	15,2371			
Muestra con DA	91	96,1271	14,5529	14,653	187	0,000
Total	189	100,4359	15,4437			

En cuanto al índice de homogeneidad, es decir, la correlación que mantiene cada ítem con el total de la puntuación obtenida, hemos calculado las correlaciones para cada ítem y también el índice promedio, siendo los resultados obtenidos especialmente positivos.

En este sentido, 55 de las correlaciones efectuadas son significativas al 0,01 y 7 lo son al 0,05; 15 de las correlaciones son positivas aunque no llegan a ser significativas, y las 3 restantes presentan una correlación negativa, que no supera en ningún caso el valor -0,098.

Por otra, parte, podemos comprobar que las puntuaciones de la muestra en su conjunto no se adecuan totalmente a la curva normal. Esto podría ser debido a que el grupo de sujetos no haya sido lo suficientemente grande. Por su parte, los índices de asimetría y curtosis expresan una distribución de las puntuaciones ligeramente asimétrica y con un valor negativo, indicando que hay un porcentaje mayor de sujetos que han obtenido valoraciones positivas y puntuaciones altas en la escala.

Podemos también observar que la media obtenida por las chicas es notablemente superior a la que muestran los chicos, presentando diferencias estadísticamente significativas ($t = 2,731$; sig. 0,008, para 96 grados de libertad). El cálculo de la prueba chi-cuadrado para comprobar si existen diferencias significativas entre el número de niñas y de niños que componen la muestra, arroja en cambio resultados no significativos (chi-cuadrado = 1,469; sig. 0,225, para 1 grado de libertad). También en cuanto al nivel educativo hemos encontrado diferencias significativas. Estas diferencias pueden resultar en principio bastante sorprendentes, ya que a medida que subimos en nivel educativo, la puntuación media de los grupos es menor. La explicación a estos resultados podría enfocarse desde dos perspectivas distintas: por una parte, es posible que el alumnado de mayor edad y nivel educativo se haya expresado con una mayor conciencia de sus propias dificultades y limitaciones, al tiempo que con mayor grado de sinceridad; por otra parte, también es posible admitir que los niveles de adaptación y ajuste social y personal relacionados con el ámbito de la lectura, sean menores en un buen número de alumnos a medida que avanzan en el sistema educativo. Aunque se encuentran diferencias significativas cuando realizamos una comparación que contemple conjuntamente todos los subgrupos, los contrastes múltiples realizados posteriormente, contrastes que se han realizado mediante la prueba T2 de Tamhane, una prueba conservadora de comparaciones por parejas basada en la prueba t y especialmente adecuada cuando no existe homogeneidad de varianzas, nos han revelado que estas diferencias no llegan a ser estadísticamente significativas para ninguno de los contrastes efectuados.

Por otra parte, en la validación de la escala APSL, el análisis de los ítems para determinar la facilidad o dificultad de los mismos podría tener un sentido distinto en función del objeto de estudio de la prueba. En efecto, mediante la aplicación de la escala no se evalúa el conocimiento por parte del sujeto acerca de un contenido de tipo conceptual o procedimental, sino que se trata de acceder a la valoración del conocimiento y la percepción del sujeto sobre contenidos de tipo actitudinal y que hacen referencia a uno mismo. Por tanto, en nuestro análisis de la escala, lo que sí será de interés es determinar qué ítems de respuesta con valoración positiva muestran un mayor consenso en cuanto al acuerdo o el desacuerdo en las respuestas y, por consiguiente, son ítems en los que su contenido denota una menor dificultad para el alumnado en relación con el ajuste personal-social en la lectura. Por el contrario, aquellos ítems de respuesta con valoración positiva en los que haya un menor consenso en cuanto al acuerdo o desacuerdo mostrado en las respuestas del alumnado, serían ítems cuyo contenido reflejaría una mayor dificultad en relación con los procesos de ajuste en la lectura. En la

siguiente tabla podemos observar los índices de dificultad calculados para cada uno de los ítems de la escala, así como los índices de homogeneidad.

Tabla 17. Índices de Dificultad (ID) e Índices de Homogeneidad (IH) para cada uno de los ítems de la escala APSL. (Los índices de dificultad están expresados en % de alumnos que fallan el ítem).

Dimensiones	ítems	ID	IH	Dimensiones	ítems	ID	IH
D1. Conocimiento de las propias limitaciones cognitivas actuales, potencialidades y posibilidades de aprendizaje, dificultades y vías de solución.	1	53,06	0,026	D 2. Conocimiento de las propias expectativas, autoconcepto académico y estilo atribucional	6	48,98	0,490**
	2	41,84	0,106		42	6,12	0,246*
	17	28,57	0,429**		52	6,12	0,337**
	22	44,90	0,252*		4	26,53	0,390**
	30	18,37	0,332**		15	17,35	0,257*
	39	39,80	0,405**		26	16,33	0,396**
	41	58,16	0,440**		46	28,57	0,320**
	53	35,71	0,620**		59	32,65	0,517**
	72	34,69	0,509**		66	15,31	0,540**
D 3. Conocimiento de la propia actitud ante la lectura: emociones, sentimiento de rechazo vs satisfacción y gusto por la lectura, tolerancia a la frustración, persistencia ante la dificultad, autonomía vs dependencia.	76	36,73	0,467**	D 4. Conocimiento de los propios intereses y motivaciones	19	59,18	0,472**
	8	8,16	0,302**		38	74,49	0,168
	11	25,51	-0,054		75	31,63	0,328**
	14	18,37	0,356**		10	35,71	0,477**
	18	20,41	0,210*		13	33,67	0,526**
	28	29,59	0,295**		32	25,51	0,416**
	34	8,16	0,267**		45	37,76	0,566**
	36	8,16	0,281**		57	41,84	0,516**
	40	15,31	0,488**		68	52,04	0,360**
	50	58,16	0,415**	70	48,98	0,504**	
	55	13,27	0,290**	D 6. Percepción de la relevancia del tema: la lectura, para que sirve...	54	6,12	0,143
	73	13,27	0,282**		60	11,22	0,360**
	7	37,76	0,348**		79	6,12	0,111
	24	55,10	0,321**		12	13,27	0,501**
	51	17,35	0,340**		47	17,35	0,358**
69	17,35	0,114	D 7. Conoc. de las propias estrategias de relación...	31	31,63	0,140	
71	61,22	0,414**		63	55,10	0,087	
78	13,27	0,145		64	30,61	0,481**	
80	61,22	0,287**	D 8. Conoc. de las propias estrategias de control y regulación consciente del proceso de lectura (planificación, supervisión y autoevaluación).	5	21,43	0,281**	
D 5. Percepción del apoyo que ofrece el contexto escolar y familiar, y de las condiciones en las que se desenvuelve.	21	52,04		0,122	16	32,65	0,024
	37	56,12		0,211*	29	14,29	0,351**
	58	36,73		0,307**	49	13,27	0,313**
	65	39,80		0,208*	9	41,84	0,471**
	67	44,90		-0,098	20	36,73	0,651**
	74	52,04		0,122	23	23,47	0,358**
	3	60,20		-0,078	25	39,80	0,251*
	27	18,37		0,191	33	24,49	0,518**
	35	24,49	0,409**	D 9. Recoconoc. de la generalización...	43	31,63	0,426**
	44	8,16	0,163		61	36,73	0,155
48	10,20	0,321**	(IH) ** La correlación es significativa al 0,01 * La correlación es significativa al 0,05				
56	22,45	0,301**					
62	13,27	0,377**					
77	11,22	0,267**					

Como apuntábamos anteriormente, los niveles de dificultad han sido relativamente bajos, con 32 ítems cuyos ID están por debajo del 25% y tan sólo 4 que se encuentran por encima del 60%. Ya veíamos en la tabla anterior que el ID promedio había sido de 30,63% y la fiabilidad para la escala de 0,8831. Un análisis que nos ha revelado cómo

se comportaría este índice de fiabilidad en el caso de eliminar de la escala algunos de los ítems, nos muestra que, en general, los índices no varían de forma significativa, superando por muy poco en trece ocasiones el nivel citado. La exploración de los ítems cuya hipotética supresión causaría un incremento de este indicador, revela que se trata de aquellos ítems de respuesta que han resultado, con respecto al promedio de dificultad, excesivamente difíciles —o que al menos han superado este promedio de dificultad— y que, además, no han correlacionado significativamente con la puntuación total de la escala; 11 ítems se encuentran en este caso. Los dos restantes han superado también el promedio de dificultad, aunque en estos casos sí han obtenido correlación significativa con la puntuación total. En ningún caso los ID obtenidos tras la eliminación de cada uno de estos ítems, han supuesto un incremento que resultara significativo con respecto al obtenido con todos los ítems contemplados en la escala; así mismo, ningún ID global obtenido ha superado el índice alfa estandarizado que se situaba en 88,74. Por otra parte, aunque en relación con lo anteriormente expuesto, ya comentamos que los IH eran excelentes, mostrando en general que la gran mayoría de los ítems contribuye de forma significativa a medir lo mismo que se está valorando con el conjunto de la escala.

Otro aspecto de interés consistía en comprobar cómo se relacionaban entre sí y con la puntuación total las distintas dimensiones contempladas en la escala. Nos interesaba conocer si las distintas dimensiones contribuían de forma conjunta a explicar los resultados y, por consiguiente, un mayor ajuste personal-social en los sujetos. Para ello hemos efectuado el análisis de una matriz de correlaciones entre las puntuaciones obtenidas en las distintas dimensiones de las que se compone la escala. Hemos calculado así mismo la correlación para la suma de las dimensiones 8 y 9, por ser de interés en el estudio, ya que se trata de los dos grupos de cuestiones más directamente relacionados con el metaconocimiento estratégico. Hemos realizado también los cálculos de la correlación con la puntuación total obtenida. Los resultados se muestran en la tabla 18.

Podemos observar que el número de correlaciones significativas es muy alto; 36 muestran una correlación altamente significativa (con una significación asociada inferior al nivel 0,01) y 10 muestran correlación significativa al nivel de 0,05. Las restantes 9, aunque son positivas, no llegan a ser significativas estadísticamente. El índice promedio de correlación es de 0,3944. Por otra parte, la puntuación total obtenida en la escala correlaciona muy significativamente con todas las dimensiones consideradas. Estos datos, unidos a que la media de las puntuaciones para cada una de las dimensiones no difiere de forma significativa, apuntan hacia la existencia de una interrelación consistente entre el contenido evaluado por las distintas dimensiones observadas en la escala, lo que a su vez informa sobre la unidad que presenta la prueba en cuanto a estos bloques considerados. En este sentido, la calidad del ajuste personal-social manifestado para cada una de estas dimensiones se reflejaría de forma significativa en la calidad del ajuste para otras dimensiones o, dicho de otro modo, la puntuación obtenida en cada una de las dimensiones explicaría en parte la puntuación en las otras. Por otra parte, la exploración de los datos también nos hace destacar las dimensiones 7 y 9 como aquellas en las que se han obtenido índices más bajos de relación con el resto de bloques o dimensiones; de las 9 correlaciones no significativas que encontramos en la matriz, 6 están relacionadas con la dimensión 7, y 3 lo están con la 9; entre el resto de dimensiones no existe ningún contraste no significativo. Un dato que permite explicar de otro modo este hecho lo obtenemos si analizamos los índices de

DIMENSIONES DE LA ESCALA APSL

- D1 Conocimiento de las propias limitaciones cognitivas actuales, potencialidades y posibilidades de aprendizaje, dificultades y vías de solución.
 D2 Conocimiento de las propias expectativas, autoconcepto académico y estilo atribucional.
 D3 Conocimiento de la propia actitud ante la lectura: emociones, sentimiento de rechazo vs satisfacción y gusto por la lectura, tolerancia a la frustración, persistencia ante la dificultad, autonomía vs dependencia.
 D4 Conocimiento de los propios intereses y motivaciones.
 D5 Percepción del apoyo que ofrece el contexto escolar y familiar, y de las condiciones en las que se desenvuelve.
 D6 Percepción de la relevancia del tema: la lectura, para que sirve, cual es su importancia.
 D7 Conocimiento de las propias estrategias de relación con los demás en tareas de lectura.
 D8 Conocimiento de las propias estrategias de control y regulación consciente del proceso de lectura (planificación, supervisión y autoevaluación).
 D9 Reconocimiento de la capacidad de generalización de estos conocimientos o estrategias.

Tabla 18. Matriz de correlaciones entre las distintas dimensiones contempladas en la escala APSL

	D 1	D 2	D 3	D 4	D 5	D 6	D 7	D 8	D 9	D 8 y 9	Punt. total
D 1	1,000	0,611**	0,474**	0,392**	0,203*	0,257*	0,176	0,504**	0,251*	0,538**	0,727**
D 2	0,611**	1,000	0,496**	0,384**	0,258*	0,366**	0,117	0,414**	0,234*	0,452**	0,714**
D 3	0,474**	0,496**	1,000	0,437**	0,214*	0,451**	0,058	0,612**	0,295**	0,649**	0,760**
D 4	0,392**	0,384**	0,437**	1,000	0,286**	0,361**	0,343**	0,523**	0,213*	0,541**	0,744**
D 5	0,203*	0,258*	0,214*	0,286**	1,000	0,319**	0,249*	0,235*	0,148	0,261**	0,517**
D 6	0,257*	0,366**	0,451**	0,361**	0,319**	1,000	0,227*	0,426**	0,175	0,442**	0,575**
D 7	0,176	0,117	0,058	0,343**	0,249*	0,227*	1,000	0,118	0,048	0,122	0,344**
D 8	0,504**	0,414**	0,612**	0,523**	0,235*	0,426**	0,118	1,000	0,129	0,936**	0,744**
D 9	0,251*	0,234*	0,295**	0,213*	0,148	0,175	0,048	0,129	1,000	0,470**	0,364**
D 8 y 9	0,538**	0,452**	0,649**	0,541**	0,261**	0,442**	0,122	0,936**	0,470**	1,000	0,791**
Puntuación total	0,727**	0,714**	0,760**	0,744**	0,517**	0,575**	0,344**	0,744**	0,364**	0,791**	1,000

** La correlación es significativa al nivel 0,01 (bilateral) * La correlación es significativa al nivel 0,05 (bilateral)

homogeneidad de sus ítems; observamos que sólo uno de los tres que conforman el bloque 7 presenta correlación significativa con la puntuación total de la escala (el ítem 64); en cuanto al bloque 9, también uno de sus dos ítems presenta correlación significativa con el total. Es posible que haya incidido en estos resultados el hecho de que estas dos dimensiones contaban con un número de ítems significativamente menor que las demás.

Hemos calculado el índice alfa de Cronbach para la escala en función de sus dimensiones y hemos procedido a eliminar cada una de las dimensiones para comprobar si la fiabilidad de la prueba experimentaba cambios cuando se prescindía de la influencia de uno de los bloques. En efecto, aunque el índice de fiabilidad era bastante alto (0,8773), hemos comprobado que éste se incrementaba (superando 0,8832) hasta rebasar el índice alfa estandarizado (0,8775) cuando eliminábamos tanto la dimensión 7 como la 9. Este dato podría ser útil en un próximo trabajo, en el que quizás la escala podría reducir el número de dimensiones, incluyendo algunos de los ítems

pertenecientes a los bloques que han mostrado una menor relación con el contenido del resto de la prueba, en aquellas dimensiones con contenido más afín.

Las puntuaciones de la escala APSL han sido transformadas de forma similar a las que se obtienen de la prueba de inteligencia utilizada en nuestro estudio. La conversión de las puntuaciones directas en puntuaciones típicas tiene sentido en relación a facilitar las comparaciones con otras pruebas similares cuyas puntuaciones hayan sido transformadas de igual forma. Así mismo, la tipificación se realiza con el objetivo — más adecuado a nuestro caso— de encuadrar una determinada puntuación con respecto a una muestra de referencia. De este modo, la puntuación así obtenida puede interpretarse, no solo con relación a sí misma, sino también en relación al grupo de sujetos de referencia que ha realizado la prueba. Como es sabido, las puntuaciones típicas expresan el número de desviaciones típicas que una determinada puntuación se separa de la media poblacional de referencia (Botella, León, San Martín y Barriopedro (2001) y cumplen los objetivos que hemos expresado más arriba. En cambio, estas puntuaciones tienen el inconveniente de que al tener como media cero y como desviación típica uno, una parte importante de las puntuaciones resultantes suelen ser negativas y casi todas son decimales. Esto resulta en ocasiones difícil de manejar y hace incómodo su tratamiento y presentación. Una solución que suele aplicarse para hacer más fácil y cómoda la interpretación y presentación de los resultados, consiste en transformar las puntuaciones típicas en una *escala derivada*. Las puntuaciones resultantes de esta transformación conservan todas las propiedades expresadas anteriormente y son equivalentes a las originales. La obtención de esta escala derivada se realiza multiplicando cada puntuación típica por una constante y sumándole posteriormente otra constante. La desviación típica de estas nuevas puntuaciones será igual a la primera constante aplicada y la media será la segunda constante. De esta forma, pueden construirse diferentes escalas derivadas que seguirán manteniendo la equivalencia con las puntuaciones originales. En nuestro caso, como ya apuntábamos en páginas anteriores, decidimos construir y aplicar una escala derivada con media 100 y desviación típica 16, por lo que todas las puntuaciones manejadas se corresponden con esta transformación.

La validación de expertos

Previamente a su aplicación dentro de la fase experimental del estudio, la escala APSL fue propuesta para su validación a través del análisis efectuado por cuatro profesionales del ámbito educativo (una orientadora de un IES, dos orientadores miembros de Equipos de Orientación Educativa para la etapa de educación primaria, y una profesora de apoyo a la integración), todos ellos con formación en psicología o pedagogía y experiencia de varios años en el campo de la educación y de la atención a las necesidades educativas especiales. Aunque en un principio se esperaba contar con un número similar de profesionales del ámbito académico, finalmente esto no fue posible debido, por una parte, a las dificultades prácticas que conllevó la búsqueda, encuentro y posterior colaboración de otros investigadores que cumplieran el perfil requerido para poder evaluar las pruebas elaboradas, y, por otra, a la relativa urgencia que acompaña en ocasiones a la investigación educativa, ligada necesariamente a los ritmos escolares. La selección de los expertos se realizó por tanto entre aquellos profesionales que habían mostrado su interés y colaboración en anteriores estudios realizados por nosotros.

El calendario aproximado para esta validación constó de los siguientes momentos: (a) elaboración del material que sería distribuido entre los colaboradores para realizar la valoración; (b) selección final de expertos; (c) proceso de validación; (d) recogida y análisis de los datos; (e) conclusiones y modificaciones en función del proceso de validación; y (f) versión definitiva de la escala.

En cuanto al material utilizado para el proceso de validación (anexo VIII), éste ha estado compuesto por:

- Una plantilla que recogía la totalidad de los ítems de la escala, clasificados a su vez en función de las distintas dimensiones contempladas, así como una escala de valoración cuantitativa donde los expertos debían marcar de 1 a 4 en función de la adecuación de cada uno de los ítems a la evaluación de las distintas dimensiones contempladas en la escala. En la plantilla se recogía así mismo una gradación cualitativa que permitía comprender y realizar la valoración, así como interpretar posteriormente las marcaciones efectuadas. Esta gradación contenía cuatro niveles:

(1) *Nivel bajo o muy bajo*. El nivel de adecuación del ítem para la evaluación de las dimensiones descritas es bajo o muy bajo.

(2) *Nivel medio-bajo*. El nivel de adecuación no es satisfactorio. El contenido del ítem se muestra claramente insuficiente para evaluar la dimensión contemplada.

(3) *Nivel medio-alto*. El nivel de adecuación del ítem para la evaluación de las dimensiones contempladas es satisfactorio. El contenido del ítem evalúa dichas dimensiones.

(4) *Nivel alto o muy alto*. El nivel de adecuación es notable. El contenido del ítem evalúa de forma adecuada o muy adecuada las dimensiones descritas.

Esta plantilla también recogía una escala de valoración cuantitativa donde los expertos debían valorar globalmente la escala, otorgando así mismo de 1 a 4 puntos en función de su adecuación para evaluar la variable socio-emocional con relación a la lectura.

- Un cuestionario de preguntas abiertas sobre diversos aspectos de la escala, donde los expertos debían realizar una valoración cualitativa acerca de estos aspectos y donde podían aportar todas las sugerencias que consideraran oportunas sobre los elementos integrantes de la escala en su conjunto para la mejora del instrumento y su adecuación a los objetivos del estudio. Las cuestiones sobre las que debía realizarse esta valoración cualitativa eran:

- * Sobre las dimensiones teóricas contempladas en la escala
- * Sobre el sistema de marcación utilizado
- * Sobre el sistema de puntuación
- * Sobre el sistema de anulación de respuestas
- * Sobre la expresión/redacción de los ítems y su adecuación al alumnado
- * Otros

La metodología que se ha empleado para el análisis de esta valoración por parte de los expertos seleccionados, ha sido esencialmente cualitativa, ya que aunque se recogen datos proporcionados por un instrumento que incluye una escala de respuesta numérica (de 1 a 4 puntos) y esto permite un cierto grado de análisis cuantitativo, la lectura e interpretación que podemos realizar de esos datos no puede dejar de centrarse en el significado que posee cada categoría designada. A continuación se recogen algunos de los datos y conclusiones que podemos extraer de la valoración realizada por nuestro grupo de expertos.

En cuanto a los datos cuantitativos, en la siguiente tabla podemos revisar una síntesis de las valoraciones efectuadas. La puntuación expresada en las columnas correspondientes a los expertos responde al promedio efectuado entre sus valoraciones de la adecuación de cada uno de los ítems contemplados en la prueba.

Tabla 19. Valoraciones cuantitativas del contenido de los ítems de la escala en función de las dimensiones contempladas (E = Experto; ---- = dimensión no valorada por el experto. Para el cálculo del promedio se ha otorgado un punto a esta dimensión)

Dimensiones de la escala		E 1	E 2	E 3	E 4	Total
Conocimiento de uno mismo y de los demás	D1. Conocimiento de las propias limitaciones...	3,60	3,70	3,70	3,70	3,68
	D2. Conocimiento de las propias expectativas...	3,44	3,56	4,00	3,67	3,67
	D3. Conocimiento de la propia actitud...	3,00	3,67	3,78	3,72	3,54
	D4. Conocimiento de los propios intereses...	3,50	3,60	3,80	3,30	3,55
	D5. Percepción del apoyo que ofrece el contexto...	3,14	4,00	3,57	3,50	3,55
	D6. Percepción de la relevancia del tema...	3,00	4,00	4,00	3,80	3,70
Metaconocimiento estratégico	D7. Conocimiento de estrategias de relación...	3,00	3,33	4,00	2,67	3,25
	D8. Conocimiento de estrategias de control...	3,67	3,67	----	3,67	3,00
	D9. Reconocimiento de la generalización...	4,00	4,00	----	3,50	3,13
Promedio de las valoraciones		3,37	3,73	3,21	3,51	3,46
Valoración global de la escala		3	3	4,00	4,00	3,50

El análisis exploratorio de estos datos nos revela que el conjunto de las valoraciones efectuadas por los expertos ha sido más que satisfactorio, situándose el promedio de estas valoraciones por encima del nivel catalogado como medio-alto. A tenor del contenido expresado para cada una de las categorías designadas, el nivel general de adecuación de los ítems para evaluar las dimensiones contempladas ha sido considerado como satisfactorio y por tanto, a juicio del grupo de expertos consultado, el contenido de los ítems evalúa dichas dimensiones. Esta valoración positiva no ha sido debida al peso desproporcionado que haya podido tener alguno de los análisis efectuados. El 91,67% de las valoraciones sobrepasa el nivel medio-alto de puntuación (3) y el 72,22% está por encima de 3,50; todas las valoraciones promedio de los expertos sobrepasan el nivel medio-alto, así como todas las valoraciones promedio de cada una de las dimensiones. Por otra parte, el promedio total de todas las valoraciones se acerca mucho al promedio total de la valoración global de la escala efectuada por los expertos.

Por lo que respecta a la valoración cualitativa realizada por los expertos a través de las preguntas abiertas del cuestionario, hemos realizado un análisis del contenido de cada una de las respuestas y valoraciones efectuadas en función de las cuestiones planteadas.

* Dimensiones teóricas contempladas en la escala

En general fueron consideradas adecuadas. El esquema propuesto se valoró como completo y claro, aunque se expresó la conveniencia de establecer diferencias más nítidas entre el contenido de unas dimensiones y otras, ya que en ocasiones podría ser fácil confundir algunos ítems con respecto a la dimensión que evalúan. También se expresó en este sentido, que algunas respuestas dadas a determinados ítems, podrían verse afectadas por elementos y procesos recogidos en otras dimensiones. Por ejemplo, las respuestas a una de las cuestiones relacionadas con la percepción del apoyo que ofrece el contexto escolar y familiar, podrían verse condicionadas por el autoconcepto y las propias expectativas.

En realidad, el hecho de que las respuestas dadas a ítems pertenecientes a una determinada dimensión, se vean condicionadas o afectadas por procesos relacionados con el ajuste personal-social de otro bloque contemplado, no deja de ser lo esperable al evaluar procesos de esta índole, ya que las distintas dimensiones contempladas tratan de acercarse a la valoración de un proceso complejo que aglutina a su vez a una serie de sub-procesos y elementos también complejos. El intento necesario de evaluar de forma «fina» o aislada uno de estos elementos —como el autoconcepto— no debe hacernos olvidar que cualquier respuesta del sujeto sobre esta percepción estará sin duda condicionada en mayor o menor medida por las experiencias relacionadas con el apoyo que ofrece el contexto escolar o familiar, así como por las condiciones en las que este apoyo se desenvuelve; a su vez, esta relación la podemos encontrar en sentido inverso.

* Sobre el sistema de marcación utilizado

El análisis en este punto se centró en las dificultades que para algunos alumnos podría suponer la utilización durante la prueba de los conceptos *acuerdo* y *desacuerdo*. En este sentido, tanto la aplicación experimental de la prueba que algunos miembros del grupo de expertos hicieron para completar su análisis, como la posterior aplicación piloto, confirmaron que, en algunos casos, los alumnos necesitaban transformar en *sí* el *acuerdo* y en *no* el *desacuerdo*.

* El sistema de puntuación y el sistema de anulación de respuestas

El análisis del sistema de puntuación ha sido realizado por algunos expertos en consonancia con el análisis del sistema de anulación de respuestas; por este motivo, procedemos al análisis conjunto de estas dos cuestiones. En este sentido, se expresó por parte de uno de los profesionales que la corrección de la prueba con el sistema de anulación de respuestas se hacía lenta y compleja, y que tal vez las posibles incoherencias manifestadas por el alumnado en sus respuestas podrían verse ya reflejadas en la puntuación debido a un efecto de compensación. Por otra parte, se aludía, igualmente en cuanto al sistema de anulación de respuestas, que podría adolecer de suficiente validez el hecho de anular dos respuestas dadas en función de diferentes dimensiones (se hacía referencia a ciertos ítems de un determinado bloque, para los

cuales se había propuesto originalmente una anulación de la respuesta en caso de que se diera una determinada respuesta ante un ítem que pertenecía a otro bloque distinto).

* Sobre la expresión/redacción de los ítems y su adecuación al alumnado

Se hizo alusión a las posibles dificultades que podrían experimentar algunos alumnos con problemas de aprendizaje a la hora de comprender los ítems. Concretamente, se hizo referencia a los ítems que estaban redactados en negativo como un factor clave que podría dificultar significativamente la comprensión de las cuestiones. Por otra parte, se apuntó la posibilidad de ofrecer apoyo al alumnado en la comprensión de los ítems y, así mismo, dividir la tarea en más de una sesión.

* Otras cuestiones planteadas por los expertos

Uno de los profesionales había dejado sin valorar cuantitativamente las dimensiones relacionadas con el metaconocimiento estratégico. En su análisis cualitativo, argumentó que el contenido de estas dimensiones podría estar valorando estrategias metacognitivas y no conocimiento metacognitivo sobre las estrategias de lectura. También hizo referencia a que la información que podría obtenerse de los ítems pertenecientes a la dimensión 9, *Reconocimiento de la capacidad de generalización de las estrategias*, aunque relevante para la valoración de la comprensión, no lo sería para la valoración del ajuste personal-social. Por otra parte, también se hizo alusión por uno de los expertos a la posible reducción del número de ítems de la escala.

El análisis conjunto posterior realizado sobre el contenido de la valoración efectuada por los expertos, nos permitió revisar minuciosamente los ítems que trataban de valorar el metaconocimiento estratégico. Este análisis confirmó la estructura y el contenido propuesto originalmente para las dos dimensiones que recogían estos ítems. En este sentido, se consideró que la redacción de los ítems, así como su contenido, remitían necesariamente a la reflexión por parte del sujeto acerca de sus propias estrategias de comprensión, así como acerca de la posibilidad de generalización de esas mismas estrategias. Por otra parte, el número de ítems, que había sido reducido ya previamente a la validación de expertos y la aplicación piloto, fue considerado razonablemente adecuado en función del objeto de la tarea y de las dimensiones teóricas contempladas en la valoración del ajuste personal-social. Ya hemos hecho alusión en páginas anteriores, tras el análisis de la aplicación piloto, que algunos de los ítems contenidos en las dimensiones que cuentan con un reducido número de éstos, podrían ser incluidos en otras dimensiones en posteriores versiones de la escala. Esta opción, tal y como reflejan algunos de los análisis efectuados, podría incrementar la validez interna y la fiabilidad de la escala, aunque no supondría en principio reducción del número de ítems contemplado.

Repercusiones de la validación de expertos para la versión final de la escala

Como apuntábamos unas líneas más arriba, la valoración efectuada por el grupo de expertos fue analizada y discutida de forma individual y colectiva con éstos. De este análisis, valorado conjuntamente con la posterior aplicación piloto de la escala, se desprendieron una serie de propuestas de cambios significativos en el contenido y en los sistemas de puntuación de la escala que contribuirían decisivamente a la versión final de la misma.

Por una parte, quizás la repercusión más importante de esta valoración de expertos fue la eliminación del sistema de anulación de respuestas. Tal y como apuntábamos cuando procedíamos a describir la escala, el análisis de los primeros resultados de la aplicación piloto —unido a los comentarios y a la propia experiencia de aplicación experimental de uno de los expertos— nos hizo ver la ineficacia del sistema de anulación de respuestas propuesto originalmente. El cómputo de las respuestas adecuadas en función de su valoración positiva asignada bastaba para establecer las diferencias que nos interesaban entre los sujetos participantes, produciéndose un efecto de compensación en la puntuación final obtenida con respecto a las respuestas que podrían anularse en función de su contenido. La eliminación del sistema de anulación supuso en la práctica una mayor agilidad en la corrección de la prueba, al tiempo que la salvación de complejos escollos teóricos derivados de la consideración de la posible anulación de dos o más respuestas dadas a ítems pertenecientes a dimensiones teóricas distintas.

Por otra parte, otro efecto significativo de la validación de expertos fue la revisión y corrección de algunos de los ítems que estaban redactados mediante oraciones negativas. En efecto, aunque, como fue explicado con anterioridad, la redacción de algunos ítems en negativo fue considerada necesaria para reforzar determinadas ideas o planteamientos expresados en la escala, el análisis exhaustivo de los mismos hizo comprobar que quizás eran demasiados, planteando dificultades innecesarias a los sujetos que participarían del estudio. Se procedió a una nueva y definitiva redacción, manteniendo en negativo sólo aquellos ítems cuyo contenido perdiera fuerza de haber sido expresado de otra forma.

Por último, en cuanto a la posibilidad de que el alumnado recibiera apoyo durante la realización de la prueba, éste era ya un aspecto contemplado originalmente en la escala, aunque no había sido recogido explícitamente en el material que fue dado a los expertos. Dado que el contenido de la escala APSL eran los procesos de ajuste personal-social, nos interesaba sobremanera que la prueba ofreciera una valoración precisamente acerca de ese contenido, sin que pudiera estar afectada esta valoración por la capacidad de comprensión del sujeto. En este sentido, en las instrucciones dadas a los profesionales que más tarde colaborarían en la pasación de las pruebas, se hacía explícita la posibilidad de leer en voz alta cuantas veces fuera necesario el contenido de los ítems.

Las pruebas de comprensión de textos. La prueba ECO 1 y la prueba ECOS

Se han utilizado también dos pruebas de evaluación de la comprensión de textos, la prueba *ECO 1*, pensada para su aplicación en educación primaria, y la prueba *ECOS*, elaborada para evaluar la comprensión en la etapa de educación secundaria obligatoria. Ambas pruebas han sido elaboradas por nosotros y pretenden dar respuesta a la valoración del conocimiento sobre las propias estrategias de comprensión, así como de la capacidad de comprensión global y de la utilización de habilidades relacionadas con la comprensión de los textos. Nos interesaba también especialmente, que ambas pruebas pudieran evaluar la puesta en práctica de procesos de autorregulación en la lectura. En general, el uso de estas pruebas nos permitiría principalmente:

- Poder comparar los resultados obtenidos con los que se derivaran de la aplicación del dispositivo de evaluación dinámica EDPL, que también realiza un análisis

preferente de los procesos metacognitivos en el proceso de lectura. Aunque la información que se obtiene de la aplicación de la evaluación dinámica de los procesos implicados en la lectura es esencialmente cualitativa, puede realizarse un análisis comparativo entre los resultados obtenidos en ambos instrumentos, ya que las *Hojas de registro y evaluación* para el análisis de la evaluación dinámica, permiten recoger pormenorizadamente la utilización de los procesos metacognitivos en las actividades de evaluación propuestas. Esto permitiría operativizar, en lo posible, los resultados obtenidos con el fin de poder compararlos con la pasación de la pruebas citadas. Igualmente, este mismo análisis comparativo podría ser efectuado con respecto al resto de procesos contemplados, tanto en las pruebas de comprensión como en el dispositivo EDPL.

- Poder establecer comparaciones en las fases pretest y postest entre los grupos experimental y control en función de las variables *conocimiento de estrategias de comprensión, utilización de procesos de autorregulación en la lectura y competencia en cada uno de los procesos implicados en la lectura que han sido contemplados en la prueba*, medidas con estas pruebas.
- Establecer un control de la variable *rendimiento en lectura* (nivel de comprensión de textos o competencia global en la prueba), así como posteriores comparaciones pretest-postest de la citada variable y de la variable *principal fuente de las dificultades lectoras*, medidas con estas pruebas, entre los grupos experimental y control.

Es necesario destacar que la decisión de utilizar pruebas elaboradas por nosotros mismos, tanto para la evaluación del ajuste personal-social como para la evaluación de la comprensión, obedece fundamentalmente a la consideración de cierta falta de adecuación y sensibilidad en las pruebas estandarizadas conocidas por nosotros, para poder captar los posibles cambios que pudieran producirse en el alumnado como fruto de su participación del trabajo con el dispositivo de evaluación dinámica y la intervención centrada en procesos metacognitivos. Aunque estas pruebas que hemos elaborado se caracterizan también por admitir fundamentalmente un análisis cuantitativo y por valorar de forma estática respuestas cerradas, hemos tratado de contrarrestar estos elementos, que, por otra parte, difícilmente pueden ofrecer de forma satisfactoria y completa una valoración de los logros obtenidos con procedimientos dinámicos de evaluación y con una intervención centrada en procesos de autorregulación. Así, la escala de ajuste personal-social, como tuvimos oportunidad de analizar, esta centrada en la lectura, analizando diversas dimensiones socio-emocionales relacionadas especialmente con los procesos de lectura. Por su parte, en las pruebas de comprensión se han incluido dimensiones relativas a los distintos procesos que la mayoría de los autores expertos en el tema han considerado esenciales para acceder a una valoración exhaustiva de la comprensión de textos, incluyendo los procesos metacognitivos, tanto el conocimiento sobre las propias estrategias de comprensión como la puesta en práctica de procesos de autorregulación.

Las pruebas ECO 1 y ECOS han sido diseñadas como instrumentos para operativizar la variable *nivel de comprensión de textos* en educación primaria y educación secundaria respectivamente. El diseño de la investigación que nos proponíamos realizar requería de unas pruebas relativamente fáciles de aplicar, que midieran el nivel de comprensión lectora en el alumnado objeto del estudio. Estas pruebas debían servir, así mismo, para

establecer las posibles diferencias entre los grupos control y experimental, una vez que se aplicara el tratamiento, ya que se utilizarían en las fases *pretest* y *posttest* de la investigación. Por otra parte, dado el carácter del estudio que poníamos en marcha, el contenido de estas pruebas, que debían informarnos sobre la existencia o no de mejoras en los procesos de comprensión y, como parte integrante de estos, en los procesos de control y regulación consciente de la lectura en los sujetos de la muestra, adquiriría una relevancia más que notable. En efecto, si tenemos en cuenta que el estudio ha tratado de poner de manifiesto la idoneidad de aplicar procesos de evaluación dinámica en la lectura con alumnado que presenta dificultades de aprendizaje en este ámbito, analizando de forma preferente los procesos metacognitivos implicados y actuando sobre estos mediante una intervención dirigida específicamente a la mejora de los procesos de autorregulación; las pruebas que utilizáramos debían ser sensibles, en la medida en que una prueba de carácter estático puede serlo, a la valoración de los distintos procesos implicados en la comprensión, según han sido descritos en la investigación, incluyendo los procesos metacognitivos. A continuación realizamos un análisis del contenido de cada una de estas pruebas, así como de los criterios que se han seguido para establecer su validez y fiabilidad.

Análisis del contenido de la prueba ECO 1 (evaluación de la comprensión de textos) para segundo y tercer ciclo de educación primaria

La prueba ECO 1 incluye 10 textos que fueron seleccionados o elaborados en función de los siguientes criterios:

- a) Se procuró que no fueran excesivamente largos, para evitar el cansancio. Al mismo tiempo, su estructura interna y su contenido nos debían permitir poder valorar los distintos procesos contemplados.
- b) Se buscaron temas que pudieran ser relevantes e interesantes para el alumnado al que iba dirigida la prueba.
- c) Los textos contenían algunas palabras difíciles para el alumnado de educación primaria. Este hecho servía al propósito de valorar la utilización, por parte de los alumnos, del contexto para acceder a la comprensión.
- d) Se incluyeron textos con diferente estructura y pertenecientes a diversos ámbitos relacionados con las áreas curriculares de primaria. De esta forma nos encontramos con textos literarios con estructura narrativa, un texto poético, textos con estructura expositiva y textos enumerativos. Incluimos dentro de los textos enumerativos, cuadros, tablas y esquemas, por ser elementos relevantes para el acceso a la información en situaciones académicas y también cotidianas y funcionales.
- e) Por otra parte, las áreas directamente relacionadas con los textos son: Lengua y literatura, Conocimiento del Medio y Matemáticas.

Cada uno de los textos va seguido de una serie de preguntas con opción múltiple de respuesta. Solo hay una respuesta correcta —o más adecuada en función del texto y de la pregunta formulada—. El total de preguntas de la prueba es de 43. Los alumnos debían leer un texto y a continuación cada una de las preguntas, señalando la opción de respuesta que les pareciera correcta. El tiempo estimado para la realización de esta prueba es de dos sesiones, entendiendo que cada sesión tiene una duración aproximada

de unos 45-50 minutos. Se creyó conveniente partir en dos la realización de la prueba, ya que de lo contrario podría resultar excesivamente larga y provocar cansancio en el alumnado. Por otra parte, su longitud obedece al hecho de tratar de evaluar lo más adecuadamente posible cada una de las dimensiones de la comprensión contempladas. Se estimaba oportuno, así mismo, que su aplicación tuviera lugar durante las primeras horas de la mañana.

Se optó por la fórmula *preguntas con opción múltiple de respuesta* para evitar la posible interferencia comprensión-expresión escrita, que planteaban otras fórmulas de preguntas abiertas. También se consideró a este respecto el mayor atractivo y la mayor facilidad de aplicación de pruebas como la diseñada, en función de los objetivos planteados. Las preguntas formuladas tienen cinco alternativas de respuesta, a excepción de dos preguntas, que tienen solo cuatro alternativas. La presencia de este número de alternativas, así como la formulación de las mismas, debía contribuir a minimizar de forma significativa la intervención del azar en la selección de la respuesta por parte del alumnado.

Por otra parte, la fórmula de preguntas con opción múltiple de respuesta tras la lectura de un texto, plantea otros problemas que tienen que ver con la idoneidad de este formato para evaluar la capacidad de comprensión de un texto escrito. En efecto, en ocasiones puede darse el hecho de que un alumno comprenda un texto y no comprenda las preguntas que se le formulan para valorar la comprensión del texto leído, o quizás no comprenda algunas de las alternativas de respuesta. Estos son problemas difíciles de solventar en una prueba que pretende ser relativamente fácil de aplicar y no muy costosa en tiempo. A este respecto, la formulación de las preguntas y de las opciones de respuesta ha tratado de seguir un cierto patrón por el que existen unos determinados tipos de preguntas que suelen repetirse con el objeto de facilitar al alumnado la identificación de lo que se le pide. Igualmente, se seleccionó un texto y se confeccionaron unas preguntas sobre éste con el objeto de poder servir de *práctica* al alumnado antes de comenzar la realización de la prueba. La realización de las preguntas correspondientes a este texto de práctica, con las explicaciones y aclaraciones oportunas por parte del profesor o evaluador, se consideraba de especial importancia para minimizar en lo posible el efecto de novedad que pudieran tener algunas cuestiones planteadas en la prueba para algunos alumnos, así como la dificultad que pudiera representar el tipo de formulación o de presentación de algunos de los ítems. En otro sentido hay que destacar a su favor que esta fórmula adoptada parece eficaz para valorar el nivel de comprensión en el alumnado en cuanto que permite formular preguntas que se centren especialmente en procesos o habilidades concretas que quieran ser analizadas.

El alumnado podría tener siempre presente el texto. Hemos pretendido evaluar la capacidad para extraer, integrar y comprender información de los textos propuestos. El alumno podía, por tanto, releer los textos cuando le fuera necesario. Se minimizaba de esta forma la interferencia de la variable *memoria*, que podría incidir en los resultados de la prueba, distorsionándolos. Por otra parte, su aplicación podía ser individual o colectiva. En principio, en función de la muestra seleccionada para nuestro estudio (alumnado con dificultades de aprendizaje de la lectura), sería aplicada sobre todo de manera individual o en pequeños grupos. En cualquier caso, se preveía su aplicación colectiva dentro del proceso de investigación, en la aplicación piloto que se llevaría a cabo, y que, junto con la valoración del grupo de expertos, serviría a los propósitos de mejorar y adecuar al máximo el instrumento a los objetivos del estudio. En los cuadros

que se recogen posteriormente se detalla la estructura de la prueba de comprensión ECO 1 con las distintas dimensiones contempladas. Así mismo, en el volumen de anexos (Anexo I-A) se recogen las claves de respuesta que permiten corregir la prueba.

Dimensiones de la comprensión contempladas

En función de los criterios de evaluación de la comprensión lectora contemplados en diversos trabajos (Alonso Tapia, Carriedo y Mateos, 1992; Alonso Tapia, Carriedo y González, 1992; Alonso Tapia y Carriedo, 1996; Sánchez, 1999; Moreno y García, 2000; Solè, 1992; Repetto, Tellez y Beltran, 2002); y atendiendo igualmente a los procesos implicados en la lectura y a los criterios de evaluación contemplados en nuestra propia investigación (Navarro, Lama y Molina, 2001; Navarro y Mora, 2003a), hemos considerado las siguientes dimensiones de la comprensión:

Cuadro 25. Dimensiones contempladas en la prueba de comprensión de textos ECO 1 y su contenido

Concepción sobre la lectura: Para que sirve leer, en que consiste y que implica la lectura.

Conocimiento de estrategias que pueden aplicarse: estrategias metacognitivas de autorregulación (planificación, supervisión y autoevaluación), estrategias de transferencia y generalización de lo aprendido, estrategias de integración textual y estrategias de integración del texto con los conocimientos previos.

Integración textual: acceso al significado de palabras, integración lineal de las ideas, integración de todas las ideas en un esquema, construcción de proposiciones globales, distinción de la información relevante, jerarquización y secuenciación de la información, acceso a la información contenida en cuadros, tablas y esquemas, identificación y uso de la estructura textual.

Integración del texto con los conocimientos y experiencias previas: establecimiento de relaciones, formulación de hipótesis, realización de inferencias, intención del autor, revisión crítica del texto desde lo ya conocido.

Supervisión del proceso de comprensión: detección de incoherencias y superación de las dificultades mediante la aplicación de estrategias de procesamiento léxico, sintáctico y semántico.

Hemos propuesto igualmente una clasificación de cada uno de los ítems en función de las dimensiones de la comprensión que se contemplan en la prueba. Lógicamente la adscripción de un ítem a uno de los procesos o dimensiones esta en función, no solo de la pregunta formulada sino también de las opciones de respuesta que se dan. Por otra parte, se podrá comprobar que algunos ítems figuran en más de un proceso. El esfuerzo por establecer distintos procesos de comprensión con el objeto de poder manejar una actividad tan compleja como la lectura y proponer criterios e indicadores de evaluación diferenciados, no nos debe hacer olvidar la interactividad del proceso en sí y la simultaneidad con la que estos procesos descritos se ponen en práctica. No obstante, para el análisis de datos, los ítems que figuran en más de una dimensión fueron adscritos finalmente a una de ellas en función de la consideración de su mejor adecuación. De esta forma, el ítem 20 de la prueba quedó encuadrado en la dimensión *Conocimiento de estrategias que pueden aplicarse*; los ítems 30, 31 y 32 fueron analizados como parte

integrante de la dimensión *Integración textual*; y los ítems 40, 41 y 42 fueron encuadrados en el bloque de *Supervisión y autoevaluación del proceso de comprensión*.

Cuadro 26. Dimensiones contempladas en ECO 1, procesos evaluados e ítems correspondientes

Dimensiones de la Comprensión	Procesos y elementos implicados en la lectura y que son evaluados con la prueba ECO 1	ítems
<i>Concepción sobre la lectura</i>	Para que sirve leer, en que consiste y que implica la lectura	26, 39
<i>Conocimiento de estrategias que pueden aplicarse</i>	Estrategias de planificación	10, 20
	Estrategias de supervisión	5, 33, 37
	Estrategias de autoevaluación	7, 12, 38
	Estrategias de Transferencia y generalización	12, 33, 43
	Estrategias de integración textual	11, 19, 24
	Estrategias de integración texto-conocimientos	38, 43
<i>Integración textual</i>	Construcción de proposiciones e integración lineal de las ideas	2, 8, 13
	Distinción de la información relevante y construcción de proposiciones globales	3, 15, 25, 40
	Jerarquización y secuenciación de la información	9, 22
	Acceso a la información contenida en cuadros, tablas y esquemas	27, 28, 29, 30, 31, 32
	Identificación y uso de la estructura textual	20, 22
<i>Integración del texto con los conocimientos y experiencias previas</i>	Establecimiento de relaciones	23, 30, 31, 32
	Formulación de hipótesis y predicción	23
	Realización de inferencias	1, 14, 16, 17, 18, 21, 30, 35, 36, 41, 42
	Jucio de la intención del autor	4, 34
<i>Supervisión y autoevaluación del proceso de comprensión</i>	Detección de incoherencias y superación de las dificultades mediante la aplicación de estrategias de procesamiento léxico, sintáctico y semántico	6, 40, 41, 42

Validez y fiabilidad de la prueba ECO 1

Aplicación piloto

Tal y como apuntábamos en páginas anteriores, la elaboración de pruebas concebidas para la evaluación de la comprensión de textos, así como del ajuste personal-social con relación a la lectura, obligaba a la aplicación piloto de esas pruebas con vistas a su

utilización con ciertas garantías en el proceso de investigación. La aplicación piloto de la prueba ECO 1 para la valoración de la comprensión de textos en primaria, al igual que ocurría con la escala de ajuste, permitiría observar y analizar, previamente a su aplicación experimental con los sujetos de la muestra, diversos aspectos relativos a su validez y fiabilidad, así como al tiempo de ejecución estimado, la dificultad o facilidad de los ítems propuestos, la inteligibilidad de las instrucciones planteadas o el interés, la actitud y la motivación con que el alumnado cumplimentaba la prueba. Esta aplicación piloto se ha llevado a cabo en el CEIP Manuel Sánchez Alonso, de Arahál (Sevilla). Los participantes han sido ocho grupos-clase completos de primaria (179 sujetos), pertenecientes a 4º (un grupo), 5º (cuatro grupos) y 6º curso (tres grupos). Es importante destacar que los grupos-clase han sido seleccionados de un centro ya participante en el estudio y en función de las posibilidades del investigador y la disponibilidad que ha mostrado el profesorado. Dentro de estos grupos, lógicamente, también hay alumnos y alumnas que presentan dificultades de aprendizaje, algunos de los cuales también han participado de nuestro estudio. Hemos optado por mantener a estos alumnos con su grupo a la hora de realizar esta prueba, ya que, como expresábamos con anterioridad, nos interesaba una muestra que, aunque no muy numerosa, fuera en algo representativa de la diversidad existente en la población escolar. En la siguiente tabla se muestra una síntesis de la información estadística básica con relación a la aplicación piloto de la prueba ECO 1.

Tabla 20. Indicadores estadísticos de la aplicación piloto de la prueba ECO 1

Indicadores	Resultados de la aplicación piloto de ECO 1
Muestra	179
Cursos	4º, 5º y 6º EP
Edades	9 a 12 años
Ítems de la prueba ECO 1	43
Índice de Dificultad medio	0,6487
Índice medio de Homogeneidad	0,3235
Puntuación media global	100,612
Error típico de la media	1,092
Desviación típica	14,617
Mediana	98,131
Varianza	213,678
Puntuación mínima	66,73
Puntuación máxima	140,75
Rango	74,01
Amplitud intercuartil	17,942

Asimetría	0,564														
Curtosis	0,268														
Puntuaciones centiles	<table> <tr><td>5</td><td>80,189</td></tr> <tr><td>10</td><td>84,674</td></tr> <tr><td>25</td><td>91,403</td></tr> <tr><td>50</td><td>98,131</td></tr> <tr><td>75</td><td>109,346</td></tr> <tr><td>90</td><td>122,803</td></tr> <tr><td>95</td><td>129,531</td></tr> </table>	5	80,189	10	84,674	25	91,403	50	98,131	75	109,346	90	122,803	95	129,531
5	80,189														
10	84,674														
25	91,403														
50	98,131														
75	109,346														
90	122,803														
95	129,531														
Prueba Kolmogorov-Smirnov de normalidad	0,088 (sig. 0,002)														
Fiabilidad (índice Alfa de Cronbach)	0,8114														
Media chicas	103,9683 (DT = 14,6839; N = 88)														
Media chicos	97,3677 (DT = 13,8752; N = 91)														
Nivel educativo															
Media 4° primaria	94,0539 (DT = 9,3541; N = 22)														
Media 5° primaria	97,4562 (DT = 12,2583; N = 83)														
Media 6° primaria	106,1030 (DT = 16,4989; N = 74)														

Nuestras primeras observaciones al examinar estos resultados se han dirigido a los índices de dificultad de la prueba y al índice promedio de dificultad. Podemos observar como éste apunta ya lo que más tarde tendremos la oportunidad de ver confirmado con los grupos experimental y control compuestos por alumnado con dificultades de aprendizaje. En efecto, expresado como el porcentaje promedio de alumnos y alumnas que yerran en sus respuestas, el dato indica que la prueba ha resultado relativamente compleja para el conjunto de participantes en esta aplicación piloto. Al igual que hicimos anteriormente con la interpretación de los resultados de la aplicación piloto de la escala de actitudes, podríamos comprobar que este índice promedio de dificultad se incrementa de forma significativa cuando acometemos el análisis de los grupos con dificultades de aprendizaje en lectura. Lógicamente, a pesar de haber resultado una prueba algo difícil para el conjunto de la muestra en primaria, las diferencias significativas con los resultados de aquellos alumnos con dificultades de aprendizaje, nos indicarían la validez de la prueba en cuanto a discernir entre los resultados obtenidos por alumnos sin dificultades y los resultados del alumnado con dificultades, de los que, a priori, podrían esperarse menores puntuaciones.

Para poder confirmar la validez de constructo de la prueba ECO 1 en el sentido apuntado anteriormente, al tiempo que su validez externa, en función de la designación previa por parte del profesorado y de los orientadores del alumnado con dificultades de aprendizaje relacionadas con la lectura, hemos realizado un análisis mediante la prueba *t* para muestras independientes con el objetivo de contrastar las puntuaciones del alumnado de la muestra de aplicación piloto (AP) con las del conjunto de alumnos y alumnas con dificultades de aprendizaje que han formado parte, bien del GE o bien del GC1. En este análisis no se ha asumido la igualdad de varianzas ya que análisis previos habían descartado la existencia de homocedasticidad entre los grupos. Los resultados muestran diferencias muy significativas a favor del grupo de sujetos que han participado de la aplicación piloto. Este aspecto confirma nuestra hipótesis y refuerza la validez de

la prueba ECO 1 por cuanto el alumnado con dificultades en el aprendizaje de la lectura obtiene significativamente una menor puntuación en una medida del rendimiento en lectura.

	N	Media	DT	Prueba t	gl.	Sig.
Muestra de AP	179	100,6127	14,6177			
Muestra con DA	35	83,5212	8,6588	9,358	77,836	0,000
Total	214	97,8174	15,1883			

Hemos de observar que en la comparación precedente, al igual que hiciéramos cuando realizábamos el contraste para la escala APSL, no se han eliminado de la muestra de AP a los alumnos con dificultades de aprendizaje; aún así, las diferencias se muestran muy significativas a favor del conjunto de alumnos que realizó la aplicación piloto de la prueba. Por otra parte, en cuanto a la correlación que mantiene cada ítem con el total de la puntuación obtenida (índice de homogeneidad), hemos calculado las correlaciones para cada ítem y también el índice promedio, siendo los resultados obtenidos muy positivos. Se han obtenido 34 correlaciones significativas al 0,01 y 2 al 0,05; 6 de las correlaciones son positivas aunque no llegan a ser significativas, y tan solo 1 presenta una correlación negativa, aunque no llega a ser significativa. Tanto el índice promedio de homogeneidad, como los índices para cada correlación (a excepción de uno), expresan que cada uno de los ítems de la prueba contribuye de manera significativa a evaluar lo mismo que el conjunto de los elementos que la conforman.

Para la prueba ECO 1 hemos podido comprobar que las puntuaciones de la muestra de la aplicación piloto en su conjunto se adecuan a una distribución normal. Por su parte, los índices de asimetría y curtosis expresan una distribución de las puntuaciones ligeramente asimétrica y con un valor positivo, indicando que hay un porcentaje importante de sujetos para los que la prueba no ha sido especialmente fácil, obteniendo puntuaciones no muy altas en los resultados globales. Podemos también observar, al igual que ocurría con los resultados del análisis de la escala APSL, que la media obtenida por las chicas es notablemente superior a la que muestran los chicos, presentando diferencias estadísticamente significativas ($t = 3,092$; sig. 0,002, para 177 grados de libertad). Hemos realizado el cálculo de la prueba chi-cuadrado para cerciorarnos de que no existen diferencias significativas entre el número de niñas y de niños que componen esta muestra de aplicación piloto; los resultados muestran la ausencia de diferencias relevantes (chi-cuadrado = 0,050; sig. 0,823, para 1 grado de libertad). En cuanto al nivel educativo, hemos encontrado diferencias significativas ($F = 10,354$; sig. 0,000). Los contrastes múltiples realizados posteriormente mediante la prueba T2 de Tamhane, nos han revelado que estas diferencias se concretan entre los cursos de cuarto y quinto, por una parte, y sexto, por otra. Los resultados son muy significativos a favor de sexto (sig. 0,000 en la comparación con cuarto; y sig. 0,001 en la comparación con quinto curso). En este sentido, los resultados de la aplicación de la prueba de Sheffé para establecer subconjuntos homogéneos, muestran un subconjunto formado por las puntuaciones correspondientes a 4º y 5º y otro con las puntuaciones de 6º de primaria.

Del análisis anterior sobre los índices de dificultad y de homogeneidad en la prueba, se derivan otros más específicos que apuntan hacia un análisis más minucioso sobre cada uno de los ítems con el objetivo último de examinar la existencia de elementos que pudieran significar un obstáculo para una mayor validez interna y fiabilidad de la

prueba en su conjunto. Este análisis suele centrarse en la excesiva dificultad o facilidad de algunos ítems, lo que lleva habitualmente a pensar que la información que aportan al conjunto de los resultados es irrelevante, dado que su poder de discriminación no es significativo. En la siguiente tabla podemos observar los resultados de este análisis. Hemos organizado los datos en función de las distintas dimensiones contempladas en la prueba.

Tabla 21. Índices de Dificultad (ID) e Índices de Homogeneidad (IH) para cada uno de los ítems de la prueba ECO 1. (Los índices de dificultad están expresados en % de alumnos que fallan el ítem).

Dimensiones contempladas	Ítems	Índice de dificultad	Índice de homogeneidad
D1. Concepción sobre la lectura	26	52,51	0,444**
	39	78,21	0,277**
D2. Conocimiento de estrategias que pueden aplicarse	5	53,07	0,376**
	7	62,01	0,327**
	10	59,78	0,399**
	11	82,68	0,433**
	12	82,68	0,004
	19	44,69	0,494**
	20	43,02	0,545**
	24	72,07	0,099
	33	46,37	0,444**
	37	73,74	0,447**
D3. Integración textual	38	72,07	0,172*
	43	86,59	0,097
	2	26,82	0,435**
	3	68,16	0,406**
	8	77,65	0,259**
	9	77,65	0,160*
	13	72,63	0,439**
	15	60,34	0,372**
	22	72,63	0,229**
	25	80,45	0,133
	27	68,72	0,408**
	28	60,34	0,426**
	29	40,22	0,490**
30	95,53	0,038	
31	80,45	-0,134	
32	86,59	0,054	
40	53,63	0,321**	
D4. Integración del texto con los conocimientos y experiencias previas	1	29,05	0,406**
	4	55,87	0,456**
	14	39,11	0,320**
	16	67,04	0,495**
	17	76,54	0,293**
	18	53,07	0,446**
	21	83,80	0,233**
	23	63,69	0,352**
	34	59,78	0,358**
	35	63,69	0,461**
D5. Supervisión y autoevaluación del proceso de comprensión	36	66,48	0,296**
	6	70,39	0,242**
	40	53,63	0,321**
	41	64,25	0,388**
	42	65,36	0,572**

(IH) ** La correlación es significativa al 0,01; * La correlación es significativa al 0,05

Tal y como apuntábamos anteriormente, los niveles de dificultad, al contrario de lo que ocurría con la escala de actitudes, han sido relativamente altos. 12 ítems han obtenido un índice promedio que se encuentra por encima del 75% y tan solo 2 ítems se encuentran por debajo del 30%. Ya veíamos en la tabla anterior que el ID promedio había sido de 64,87% y la fiabilidad para la escala de 0,8114. Los análisis encaminados a comprobar el comportamiento de este índice de fiabilidad en el caso de eliminar de la prueba algunos de los ítems, nos muestran que, en general, los índices varían poco con respecto al índice promedio, superando por muy poco el nivel citado en nueve ocasiones. La exploración de los ítems cuya hipotética supresión causaría un incremento de este indicador, revela que se trata de aquellos ítems de respuesta que han resultado, con respecto al promedio de dificultad, excesivamente difíciles —o que al menos han superado este promedio de dificultad— y que, además, no han correlacionado significativamente con la puntuación total de la escala; 7 ítems se encuentran en este caso. Los dos restantes han superado también el promedio de dificultad, aunque en estos casos sí han obtenido correlación significativa con la puntuación total al nivel 0,05. En ningún caso los ID obtenidos tras la eliminación de cada uno de estos ítems, han supuesto un incremento que resultara significativo con respecto al obtenido con todos los ítems contemplados en la prueba, aunque al superar el índice promedio considerado para el conjunto de la prueba, han superado también el índice alfa estandarizado, que se situaba en 0,7973. Mención especial merecen los ítems 30 y 31; el ítem 30 resultó ser el más difícil de la aplicación piloto, con un 95,53% de alumnos que fallaban su respuesta. Este índice, unido a que la correlación obtenida con la puntuación total es muy baja (0,038), parece mostrar una significativa falta de discriminación de este ítem con vistas a aportar información relevante para el estudio. El ítem 31 resultó ser el único que presentó una correlación negativa con la puntuación total obtenida, al tiempo que también se encuentra entre los que haría incrementar el índice de fiabilidad de la prueba si se procediera a su eliminación; su promedio de dificultad también fue bastante alto (80,45%). La decisión final de no eliminar estos dos ítems, al igual que el resto de ítems cuya supresión hacía incrementar la fiabilidad de la prueba, estuvo cimentada en dos razones fundamentalmente: por una parte, el índice de fiabilidad de la prueba sin eliminar ningún ítem era bastante alto, situándose por encima del índice alfa estandarizado, por lo que la prueba parecía otorgar suficientes garantías con vistas a poder ofrecer información diferencial en función del alumnado participante y del momento de pasación de la prueba (pretest/postest); por otra parte, los índices de homogeneidad en su conjunto también resultaron ser muy positivos, por lo que cabía la posibilidad de que, aunque hubiera un número de ítems excesivamente difíciles, éstos cumplieran una función discriminante relevante en el transcurso de la investigación. A este respecto, aunque ya finalizada la fase experimental del estudio, hemos realizado un contraste de tipo exploratorio entre los índices de dificultad hallados para la muestra de aplicación piloto y los encontrados para el grupo de alumnos y alumnas que formaron parte del GE y del GC1 en la fase pretest. Los resultados han confirmado en parte nuestras sospechas, es decir, incluso los ítems que habían resultado más difíciles para la muestra piloto, resultaron ser discriminantes con respecto a la muestra de alumnos con dificultades; esto fue así para 7 de los 9 casos a los que hacíamos referencia anteriormente, incrementándose así mismo el número de ítems total cuyo promedio de dificultad estaba por encima del 75% (35 ítems). Sin embargo, los ítems 30 y 31, a los que hacíamos especial mención más arriba, se mostraron sin ningún poder de discriminación, por cuanto el grupo de alumnos con dificultades, a pesar de obtener unos índices muy altos de dificultad para estos ítems (85,71% y 71,43%, respectivamente), promedió más bajo en estos índices que la muestra piloto. Quizás

estos últimos análisis indiquen con más fuerza la conveniencia de revisar la presencia de estos dos ítems en la prueba con vistas a futuras aplicaciones de ECO 1. Por otra parte, aunque en relación con lo anteriormente expuesto, ya comentamos que los IH eran excelentes, mostrando en general que la gran mayoría de los ítems contribuye de forma significativa a medir lo mismo que se está valorando con el conjunto de la prueba.

Un análisis que guardaba un especial interés para nosotros consistía en comprobar cómo se habían relacionado entre sí y con la puntuación total las distintas dimensiones contempladas en la prueba. Nuestro interés se centraba en conocer si estas dimensiones contribuían de forma conjunta a explicar los resultados y, por consiguiente, una mayor capacidad de comprensión lectora en los sujetos de la muestra. Para ello efectuamos el análisis de una matriz de correlaciones entre las puntuaciones obtenidas en las distintas dimensiones de las que se compone la prueba ECO 1. Hemos realizado también los cálculos de la correlación con la puntuación total obtenida. Los resultados se muestran en la siguiente tabla.

DIMENSIONES CONTEMPLADAS EN LA PRUEBA ECO 1

- D1 Concepción sobre la lectura.
- D2 Metaconocimiento de estrategias que pueden aplicarse.
- D3 Integración textual.
- D4 Integración del texto con los conocimientos y experiencias previas.
- D5 Supervisión y autoevaluación del proceso de comprensión.

Tabla 22. Matriz de correlaciones entre las distintas dimensiones contempladas en la prueba ECO 1

	D 1	D 2	D 3	D 4	D 5	Punt. Total
D1	1,000	0,318**	0,248**	0,453**	0,252**	0,496**
D2	0,318**	1,000	0,530**	0,553**	0,392**	0,809**
D3	0,248**	0,530**	1,000	0,593**	0,383**	0,806**
D4	0,453**	0,553**	0,593**	1,000	0,466**	0,859**
D5	0,252**	0,392**	0,383**	0,466**	1,000	0,626**
Puntuación total	0,496**	0,809**	0,806**	0,859**	0,626**	1,000
** La correlación es significativa al 0,01						

La matriz de correlaciones muestra una fuerte conexión entre las puntuaciones obtenidas en las distintas dimensiones contempladas en la prueba y, por tanto, entre el desempeño que han tenido los sujetos en cada uno de estos bloques; así mismo, podemos observar que todas las correlaciones son muy significativas al nivel 0,01, aunque la intensidad de las correlaciones varíe de unas dimensiones a otras. El índice promedio de correlación es de 0,4188. Por otra parte, la puntuación total obtenida en la prueba correlaciona muy significativamente con todas las dimensiones consideradas, mostrando unos niveles de correlación muy altos. Estos datos, unidos a que la media de las puntuaciones para cada una de las dimensiones es muy similar, no difiriendo de forma significativa, apuntan hacia la existencia de una interrelación muy consistente entre el contenido evaluado por las distintas dimensiones observadas en la prueba, lo que a su vez informa sobre la unidad que presenta la prueba en cuanto a estos bloques

considerados. En este sentido, la calidad de la comprensión de textos manifestada para cada una de estas dimensiones se reflejaría de forma significativa en la calidad de la comprensión en otras dimensiones, así como en la puntuación total, o, dicho de otro modo, la puntuación obtenida en cada una de las dimensiones explicaría en parte la puntuación en las otras. Por lo que respecta a los ítems analizados anteriormente que guardaban una baja correlación con la puntuación total y que, así mismo, habían resultado excesivamente difíciles para el conjunto de la muestra, éstos se encuentran encuadrados en las dimensiones D2 (*Metaconocimiento estratégico*; ítems 12, 24, 38 y 43) y D3 (*Integración textual*; ítems 9, 25, 30, 31 y 32). El análisis exploratorio de la fuerza de las correlaciones de la puntuación total obtenida con cada una de estas dos dimensiones (0,809 con D2 y 0,806 con D3), así como las correlaciones igualmente muy significativas que mantienen con el resto de bloques contemplados, parece restar relevancia a la incidencia que haya podido tener la presencia de estos ítems en la versión aplicada de la prueba.

Al igual que hicimos con la escala APSL, hemos calculado el índice alfa de Cronbach para la prueba en función de sus dimensiones y hemos procedido a eliminar cada una de ellas para comprobar si la fiabilidad de la prueba experimentaba cambios cuando se prescindía de la influencia de uno de los bloques. Los resultados muestran que tan solo cuando eliminamos el primer bloque, es decir, *Concepción sobre la lectura* (aquel que tiene una menor correlación con el resto de dimensiones y también con la puntuación total obtenida), el índice de fiabilidad experimenta un ligero incremento; por otra parte, el índice obtenido para el conjunto de dimensiones es de 0,7791, lo que unido a los datos analizados y expresados anteriormente en la matriz de correlaciones, así como al interés teórico que para nosotros tenía la presencia de un bloque de contenido que evaluara precisamente la concepción sobre la lectura (en qué consiste leer, qué implica y para qué sirve aprender a leer, etc.), nos ha hecho mantener todos los bloques contemplados originalmente en la prueba. En cualquier caso, sí es de interés destacar que el bloque aludido, aún siendo aquel que menor relación ha guardado con el desempeño en la prueba, ha mostrado una correlación muy significativa ($r = 0,496$) con la puntuación total obtenida en la prueba, lo que, visto de otro modo, refuerza enormemente las relaciones entre el contenido evaluado en esta dimensión y la comprensión efectiva de los textos.

Las puntuaciones de la prueba ECO 1 también han sido transformadas de forma similar a las que se obtienen de la prueba de inteligencia utilizada en nuestro estudio. Ya exponíamos anteriormente que la conversión de las puntuaciones directas en puntuaciones típicas tiene sentido en relación a facilitar las comparaciones con otras pruebas similares cuyas puntuaciones hayan sido transformadas de igual forma. Así mismo, la tipificación, y la posterior transformación en una escala derivada, se han realizado con el objetivo de encuadrar una determinada puntuación con respecto a una muestra de referencia, facilitando al tiempo la gestión de la información analizada. La puntuación así obtenida ha podido interpretarse, no solo con relación a sí misma, sino también en relación al grupo de sujetos de referencia que ha realizado la prueba.

La validación de expertos

La aplicación de la prueba ECO 1 dentro de la fase experimental del estudio requería previamente, no solo una aplicación piloto de la misma con el objetivo de evaluar el comportamiento a distintos niveles de sujetos pertenecientes a grupos-clase completos,

sino, en buena lógica, su validación a través del análisis denominado de expertos. Este análisis fue efectuado por tres profesionales del ámbito educativo (dos orientadores miembros de Equipos de Orientación Educativa para la etapa de educación primaria, y una profesora de apoyo a la integración), todos ellos con formación en psicología o pedagogía y experiencia de varios años en el campo de la educación y de la atención a las necesidades educativas especiales. Estos tres expertos también participaron en el análisis de la escala APSL. Al igual que comentábamos entonces, aunque en un principio se esperó contar con un número similar o superior de profesionales del ámbito académico, finalmente esto no fue posible por las mismas razones expuestas con anterioridad. La selección de los expertos se realizó por tanto en idénticas condiciones y por las mismas razones argumentadas ya con relación a la escala de actitudes. En cuanto al calendario aproximado para esta validación, éste constó de los siguientes momentos: (a) elaboración del material que sería distribuido entre los colaboradores para realizar la valoración; (b) selección de expertos; (c) proceso de validación; (d) recogida y análisis de los datos; (e) conclusiones y modificaciones en función del proceso de validación; y (f) versión definitiva de la prueba.

El material utilizado para el proceso de validación (anexo VI), ha estado compuesto por:

- Una plantilla que recogía la totalidad de los ítems de la prueba, clasificados a su vez en función de las distintas dimensiones contempladas, así como una escala de valoración cuantitativa donde los expertos debían marcar de 1 a 4 en función de la adecuación de cada uno de los ítems a la evaluación de las distintas dimensiones contempladas en la prueba. En la plantilla se recogía así mismo una gradación cualitativa que permitía comprender y realizar la valoración, así como interpretar posteriormente las marcaciones efectuadas. Esta gradación contenía cuatro niveles:

(1) *Nivel bajo o muy bajo*. El nivel de adecuación del ítem para la evaluación de los procesos descritos es bajo o muy bajo.

(2) *Nivel medio-bajo*. El nivel de adecuación no es satisfactorio. El contenido del ítem se muestra claramente insuficiente para evaluar los procesos contemplados.

(3) *Nivel medio-alto*. El nivel de adecuación del ítem para la evaluación de los procesos descritos es satisfactorio. El contenido del ítem evalúa dichos procesos.

(4) *Nivel alto o muy alto*. El nivel de adecuación es notable. El contenido del ítem evalúa de forma adecuada o muy adecuada los procesos descritos.

Esta plantilla también recogía una escala de valoración cuantitativa donde los expertos debían valorar globalmente la prueba ECO 1, otorgando así mismo de 1 a 4 puntos en función de su adecuación para evaluar la comprensión de textos en segundo y tercer ciclo de educación primaria.

- Un cuestionario de preguntas abiertas sobre diversos aspectos de la prueba, donde los expertos debían realizar una valoración cualitativa acerca de estos aspectos y donde podían aportar todas las sugerencias que consideraran oportunas sobre los elementos integrantes de la prueba en su conjunto para la mejora del instrumento y su adecuación a los objetivos del estudio. Esta valoración cualitativa servía al propósito de obtener más elementos de validez

externa del contenido y de la adecuación global de la prueba a los objetivos de nuestra investigación. Las cuestiones sobre las que debía realizarse esta valoración cualitativa fueron estimadas especialmente relevantes para la consideración de la bondad de la prueba. Los expertos eran informados a través de las instrucciones recogidas en la plantilla sobre la que debían realizar su valoración, de que podrían expresar ésta en forma de dudas, comentarios críticos sobre aspectos que deberían ser mejorados, valoraciones positivas, valoraciones sobre otros aspectos no recogidos en la plantilla y que se consideraran relevantes, etc. Los elementos propuestos para su valoración eran los siguientes:

- * Los textos (adecuación y contenido, expresión, tipologías de textos y estructuras contempladas, vocabulario...) Observaciones, dudas, comentarios.
- * Las dimensiones de la comprensión contempladas.
- * La prueba contemplada en su totalidad.

El método de análisis utilizado en nuestra valoración del trabajo desarrollado por el grupo de expertos ha sido semejante al que expusimos con anterioridad en nuestro análisis de la escala de actitudes. El aspecto más relevante de este análisis es su carácter esencialmente cualitativo y su orientación a la mejora del instrumento analizado. Recogemos posteriormente algunos de los datos y conclusiones que podemos extraer de la valoración realizada por nuestro grupo de expertos. En cuanto a los datos provenientes del análisis más cuantitativo, en la siguiente tabla podemos observar una síntesis de las valoraciones efectuadas. La puntuación expresada en las columnas de los expertos responde al promedio efectuado entre sus valoraciones de la adecuación de cada uno de los ítems contemplados en la prueba.

Tabla 23. Valoraciones cuantitativas del contenido de los ítems de la prueba ECO 1 en función de los procesos contemplados (E = Experto; ---- = dimensión o aspecto no valorado por el experto. Para el cálculo del promedio total de la valoración global se ha utilizado en este caso el promedio de las valoraciones del experto que no realizó su valoración global)

Dimensiones de la comprensión contempladas	E 1	E 2	E 3	Total
D1. Concepción sobre la lectura	4,00	3,50	4,00	3,83
D2. Metaconocimiento de estrategias que pueden aplicarse	3,06	3,75	3,31	3,37
D3. Integración textual	3,82	3,82	3,47	3,70
D4. Integración del texto con los conocimientos y experiencias previas	3,61	3,78	3,72	3,70
D5. Supervisión y autoevaluación del proceso de comprensión	3,00	3,50	3,00	3,17
Promedio de las valoraciones	3,50	3,67	3,50	3,56
Valoración global de la prueba ECO 1	4	----	3,00	3,72

Los datos mostrados en la tabla descubren que, en general, la valoración efectuada por el grupo de expertos de la adecuación de los ítems en función del proceso en el que se encuadraban, ha sido muy positivo. El promedio de estas valoraciones se ha situado por encima del nivel catalogado como medio-alto. A tenor del contenido expresado para cada una de las categorías designadas, el nivel general de adecuación de los ítems para evaluar los procesos contemplados en la prueba ECO 1 ha sido considerado como satisfactorio y por tanto, a juicio del grupo de expertos consultado, el contenido de los ítems evalúa dichos procesos. Al igual que ya hicimos con relación a la escala de ajuste, es de nuestro interés conocer si esta valoración positiva ha sido debida al peso desproporcionado que haya podido tener alguno de los análisis efectuados. El análisis

de los datos nos revela que El 100% de las valoraciones iguala o sobrepasa el nivel medio-alto de puntuación (3) y el 66,67% iguala o está por encima de 3,50; todas las valoraciones promedio de los expertos sobrepasan el nivel medio-alto, así como también todas las valoraciones promedio de cada una de las dimensiones. Por otra parte, en este caso también el promedio total de todas las valoraciones se acerca mucho al promedio total de la valoración global de la escala efectuada por los expertos.

Destacar que los ítems 30 y 31 han sido valorados con la máxima puntuación por dos de los expertos y con tres puntos por el otro, obteniendo una valoración promedio de 3,67. En este sentido, parece ser que a priori estos ítems han sido positivamente valorados en función de su adecuación para evaluar la capacidad de acceder a la información contenida en cuadros, tablas y esquemas, así como la realización de inferencias y el establecimiento de relaciones; sin embargo, la puesta en práctica de la prueba mediante la aplicación piloto, como hemos podido comprobar, nos mostró la excesiva dificultad de estos ítems para el grupo de alumnos y alumnas en general. Estos dos ítems, así como otros, fueron encuadrados originalmente en varias dimensiones por estar relacionados con todas ellas, tal como se muestra en el cuadro 26; ha sido posteriormente, tal y como comentábamos en páginas anteriores, en función de las revisiones efectuadas tras la aplicación piloto y la validación de expertos, y ante la necesidad de encuadrarlos definitivamente en algún bloque a efectos del análisis de datos, cuando han sido clasificados en un determinado proceso.

En cuanto a la valoración cualitativa realizada por los expertos a través de las preguntas abiertas del cuestionario, hemos realizado también aquí un análisis del contenido de cada una de las respuestas y valoraciones efectuadas en función de las cuestiones planteadas. Al igual que hicimos cuando analizamos la escala de actitudes, exponemos las observaciones, dudas o comentarios críticos de los expertos para, posteriormente, centrarnos en aquellos aspectos que se vieron afectados por estas valoraciones en la versión final de la prueba.

* Sobre los textos (adecuación y contenido, expresión, tipologías de textos y estructuras contempladas, vocabulario...) Observaciones, dudas, comentarios.

Los comentarios y valoraciones efectuados por los tres profesionales apuntan a la posible dificultad en la comprensión de las preguntas, los textos o algunos de los ítems de respuesta, que pueden presentar algunos alumnos, máxime cuando la investigación está dirigida a la evaluación de las dificultades de aprendizaje de la lectura. En este sentido se habla de un nivel quizás elevado en algunos textos para niños y niñas de primaria; también se plantea la posibilidad de realizar pequeñas introducciones antes de algunos textos para facilitar la comprensión de la situación que se describe. Así mismo, se hace referencia a que podría reducirse el número de textos o de cuestiones de la prueba, ya que se considera que puede resultar larga para muchos alumnos. Por otra parte, se destaca positivamente la variedad de tipologías textuales recogidas.

* Sobre las dimensiones de la comprensión contempladas

Sobre este punto se destaca en general la adecuación de las distintas dimensiones contempladas, si bien se puntualiza que en algunos casos, los ítems correspondientes a algunas de las dimensiones se encuentran en la parte final de la prueba, dificultando el hecho de que muchos alumnos puedan realizar y responder esos ítems ya que

difícilmente llegarían a esa parte de la prueba. Se apunta en este caso la necesidad de «repartir» mejor los ítems de respuesta a lo largo de la prueba para que, en el caso de que un alumno se quede sin terminar, pueda haber contestado ítems de todas las dimensiones y, de esta forma, permitir una valoración más ajustada de sus dificultades. Por otra parte, se hace alusión a la posibilidad de incrementar el número de ítems de la dimensión *Concepción sobre la lectura*.

* Sobre la prueba contemplada en su totalidad

La valoración que efectúan los expertos sobre la prueba en general remite a dos aspectos: por una parte es considerada muy positivamente en cuanto a su estructura y contenido, susceptible de resultar muy útil o adecuada para valorar lo que se pretende; por otra, los expertos coinciden en que se trata de una prueba compleja, que puede resultar excesivamente difícil o «dura» para el alumnado con dificultades. En este sentido, se habla de la necesidad de apoyo individualizado para poder obtener buenos resultados. Se hace referencia, así mismo, a la mayor dificultad que pueden suponer las cuestiones realizadas, en comparación con los propios textos.

Repercusiones de la validación de expertos para la versión final de la prueba ECO 1

Al igual que realizamos con la escala APSL, la valoración efectuada por el grupo de expertos fue analizada y discutida de forma individual y colectiva con éstos. De este análisis, valorado conjuntamente con la posterior aplicación piloto de la escala, se desprendieron una serie de propuestas de cambios significativos en el contenido y en la estructura de la prueba, que contribuirían decisivamente a la versión final de la misma, bien por la incorporación de las propuestas o bien por el reforzamiento de algunas decisiones tomadas previamente, en función de las reflexiones derivadas de los análisis.

Hemos hecho alusión en el análisis efectuado sobre las valoraciones del grupo de expertos, que se apuntaba la necesidad de que el alumnado recibiera apoyo durante la prueba. Sin duda, y máxime en una investigación como la nuestra, esta opción nos parece la más idónea con vistas a extraer de la ejecución de la prueba una mayor y mejor información sobre las dificultades y la capacidad de aprendizaje del alumno con problemas en la comprensión; sin embargo, dadas las características del estudio y la necesidad en algunos casos de aplicación colectiva de la prueba, unido al hecho de que se trataba de una prueba elaborada con el objetivo de establecer referencias pretest y postest para los grupos experimental y control con respecto a su participación o no en la aplicación del dispositivo de evaluación dinámica de procesos implicados en la lectura, la posibilidad de recibir apoyo en esta prueba quedaba fuera de nuestros propósitos. Por otra parte, aunque es cierto que la prueba puede resultar larga (dos sesiones de 45/50 minutos cada una), lo cierto es que la aplicación piloto mostró que los grupos-clase (especialmente los grupos de 5º y 6º) la realizaban y terminaban sin especiales agobios de tiempo. La posibilidad de reducir textos y cuestiones se hacía difícil debido lógicamente a la intención de incluir en la prueba la totalidad, o el máximo posible de procesos implicados en la comprensión. En cualquier caso, la prueba no es más larga que otras pruebas de comprensión de textos comercializadas o propuestas en la literatura científica especializada, cuya duración es similar o superior a la prueba ECO 1.

Si bien era previsible que muchos de los sujetos con dificultades de aprendizaje participantes en el estudio, podrían no finalizar la prueba debido a su longitud o complejidad, la estructura de la misma, con diferentes textos y cuestiones para cada uno

de los textos, así como la distribución de los ítems de respuesta correspondientes a las distintas dimensiones contempladas a lo largo de la prueba, hacían que el objetivo de la aplicación de la prueba, en cuanto al establecimiento de un control pretest entre los grupos experimental y control, así como en cuanto al establecimiento de unas medidas del rendimiento en lectura que pudieran servir de referencia para evaluar los posibles cambios en el GE tras la aplicación del tratamiento, pudiera verse justificado y cumplido. De esta forma, tal y como apuntábamos en páginas anteriores, la extensión y complejidad de la prueba había sido fruto de dos consideraciones fundamentales, de una parte, la necesidad y conveniencia de incluir todos los procesos implicados en la lectura que pudieran ser evaluados en un sistema de prueba cerrada, y de otra, la posibilidad que ofrecía de cubrir y reflejar un amplio abanico de capacidades del alumnado que participaría en la investigación.

En definitiva, aunque éramos conscientes de la dificultad que planteaba la prueba, especialmente para el alumnado con dificultades de aprendizaje, optamos por no modificar su contenido ni reducir el número de textos o ítems. La prueba había sido modificada previamente a su validación por parte del grupo de expertos y las acertadas y relevantes observaciones efectuadas por éstos habían sido objeto de discusión previa. La decisión parecía acertada en función del comentario anterior relativo a la posibilidad que ofrecía una prueba extensa y que contemplara diferentes tipologías textuales y procesos a evaluar, de dar cuenta de la diversidad de capacidades relacionadas con la comprensión, así como de establecer diferencias entre la ejecución del alumnado con problemas de comprensión en el marco de unas referencias estándar ofrecidas por la ejecución de la muestra piloto, sin dificultades en su mayoría. Comprobaremos posteriormente que, a pesar de los buenos indicadores estadísticos obtenidos en cuanto a la fiabilidad de la prueba, los índices de homogeneidad o la correlación entre las distintas dimensiones de la comprensión contempladas, una hipótesis relativa a los resultados finalmente obtenidos en esta prueba, hace referencia a que un aspecto que ha podido incidir en los mismos, ha sido precisamente la excesiva dificultad que para los alumnos y las alumnas con dificultades de lectura, ha supuesto la realización de la prueba, originando tal vez un cierto efecto techo. En la discusión sobre los resultados ampliaremos nuestro análisis de estos resultados.

Análisis del contenido de la prueba ECOS (evaluación de la comprensión de textos) para educación secundaria

La prueba ECOS incluye también 10 textos que han sido seleccionados o elaborados en función de criterios similares a los que se recogían para la prueba homóloga de primaria:

- a) Se ha procurado que no fueran excesivamente largos, para evitar el cansancio. Al mismo tiempo, su estructura interna y su contenido nos permite poder valorar los distintos procesos contemplados.
- b) Se han buscado temas que pudieran ser relevantes e interesantes para el alumnado al que va dirigida la prueba.
- c) Los textos contienen algunas palabras difíciles para el alumnado de E.S.O. Este hecho sirve al propósito de valorar la utilización, por parte de los alumnos, del contexto para acceder a la comprensión.

d) Se incluyen textos con diferente estructura y pertenecientes a diversos ámbitos relacionados con las áreas curriculares en E.S.O. De esta forma nos encontramos con textos con estructura narrativa, textos con estructura expositiva, textos periodísticos y textos con estructura inductiva-instructiva. Por otra parte, las áreas directamente relacionadas son: Lengua y literatura, Ciencias Sociales y Ciencias Naturales.

e) Hemos incluido tablas y esquemas, por ser elementos relevantes para el acceso a la información en situaciones académicas y también cotidianas y funcionales.

Cada uno de los textos va seguido de una serie de preguntas con opción múltiple de respuesta. Solo hay una respuesta correcta —o más adecuada en función del texto y de la pregunta formulada—. El total de preguntas de la prueba es de 56. Los alumnos deben leer un texto y a continuación cada una de las preguntas, señalando de entre las que se ofrecen, la opción de respuesta que les parezca correcta. El tiempo estimado para la cumplimentación de esta prueba es de dos sesiones, entendiendo que cada sesión es de unos 50 minutos. Se ha creído conveniente partir en dos la realización de la prueba, ya que de lo contrario podría resultar excesivamente larga y provocar cansancio. Por otra parte, su longitud obedece al hecho de tratar de evaluar lo más adecuadamente posible cada una de las dimensiones de la comprensión contempladas. Se estima oportuno que se aplique durante las primeras horas de la mañana.

Se ha optado también para esta prueba por la fórmula *preguntas con opción múltiple de respuesta*, con la misma intención de evitar la posible interferencia comprensión-expresión escrita que planteaban otras fórmulas de preguntas abiertas. También se ha considerado a este respecto el mayor atractivo y la mayor facilidad de aplicación de pruebas como la diseñada, en función de los objetivos planteados. Las preguntas formuladas tienen cinco alternativas de respuesta, a excepción de dos preguntas, que tienen solo cuatro alternativas. La presencia de este número de alternativas, así como la formulación de las mismas, contribuye a minimizar de forma significativa la intervención del azar en la selección de la respuesta por parte del alumnado.

Como ya vimos cuando analizábamos el contenido de la prueba de comprensión elaborada para primaria, la fórmula de preguntas con opción múltiple de respuesta tras la lectura de un texto, planteaba otros problemas que tenían que ver con la idoneidad de este formato para evaluar la capacidad de comprensión de un texto escrito. Los comentarios críticos y las salvedades que hicimos en aquella ocasión, nos servirían igualmente para describir el proceso seguido con la prueba ECOS para tratar en lo posible de maximizar las ventajas de utilizar este formato de actividad y minimizar sus inconvenientes. En este sentido, la formulación de las preguntas y de las opciones de respuesta también ha tratado de seguir un cierto patrón por el que existen unos determinados tipos de preguntas que suelen repetirse con el objeto de facilitar al alumnado la identificación de lo que se le pide. Igualmente, se ha seleccionado un texto y se han confeccionado unas preguntas con el objeto de poder servir de *práctica* al alumnado antes de comenzar la realización de la prueba. La realización de las preguntas correspondientes a este texto de práctica, con las explicaciones y aclaraciones oportunas por parte del profesor o evaluador, se ha considerado de especial importancia para minimizar en lo posible el efecto de novedad que puedan tener algunas cuestiones planteadas en la prueba para algunos alumnos, así como la dificultad que pueda representar el tipo de formulación o de presentación de algunos de los ítems.

Así mismo, el alumnado podría tener siempre presente el texto. Pretendíamos evaluar la capacidad para extraer, integrar y comprender información de los textos propuestos. El alumno podría, por tanto, releer los textos cuando le fuera necesario. Se minimizaría, de esta forma, la posible interferencia de la variable *memoria*, que podría incidir de forma significativa en los resultados de la prueba. Por otra parte, su aplicación podía ser individual o colectiva. En principio, en función de la muestra seleccionada para nuestro estudio (alumnado que presentaba especiales dificultades de aprendizaje en la lectura), su aplicación se efectuaría, sobre todo, individualmente o en pequeños grupos. En cualquier caso, su aplicación colectiva, dentro del estudio, se contemplaba en el marco de la aplicación piloto que se llevaría a cabo, y que, junto con la valoración del grupo de expertos, serviría a los propósitos de mejorar y adecuar al máximo el instrumento a los objetivos del estudio. En los cuadros que se recogen posteriormente se detalla la estructura de la prueba ECOS con las distintas dimensiones contempladas, y en el Anexo II-A las claves de respuesta que permiten corregir la misma.

Dimensiones de la comprensión contempladas

En función de los criterios de evaluación de la comprensión lectora contemplados en diversos trabajos, a los que ya hacíamos alusión con relación a la prueba dirigida al alumnado de primaria (Alonso Tapia, Carriedo y Mateos, 1992; Alonso Tapia, Carriedo y González, 1992; Alonso Tapia y Carriedo, 1996; Sánchez, 1999; Moreno y García, 2000; Solè, 1992; Repetto, Tellez y Beltran, 2002); y atendiendo igualmente a los procesos implicados en la lectura y a los criterios de evaluación contemplados en nuestra propia investigación (Navarro, Lama y Molina, 2001; Navarro y Mora, 2003a), hemos considerado las siguientes dimensiones de la comprensión en la prueba ECOS:

Cuadro 27. Dimensiones contempladas en la prueba de comprensión de textos ECOS, y su contenido

Concepción sobre la lectura: Para que sirve leer, en que consiste y que implica la lectura.

Conocimiento de estrategias que pueden aplicarse: estrategias metacognitivas de autorregulación (planificación, supervisión y autoevaluación), estrategias de transferencia y generalización de lo aprendido, estrategias de integración textual y estrategias de integración del texto con los conocimientos previos.

Conocimientos previos: Conocimientos sobre el tema del texto, sobre las situaciones planteadas y sobre los esquemas de acción descritos.

Integración textual: acceso al significado de palabras, integración lineal de las ideas, integración de todas las ideas en un esquema, construcción de proposiciones globales, distinción de la información relevante, jerarquización y secuenciación de la información, acceso a la información contenida en cuadros, tablas y esquemas, identificación y uso de la estructura textual.

Integración del texto con los conocimientos y experiencias previas: establecimiento de relaciones, formulación de hipótesis, realización de inferencias, intención del autor, revisión crítica del texto desde lo ya conocido.

Supervisión del proceso de comprensión: detección de incoherencias y superación de las dificultades mediante la aplicación de estrategias de procesamiento léxico, sintáctico y semántico.

Al igual que hicimos con la prueba pensada para su aplicación en educación primaria, hemos propuesto igualmente una clasificación de cada uno de los ítems en función de las dimensiones de la comprensión que se contemplan en la prueba. Lógicamente la adscripción de un ítem a uno de los procesos o dimensiones, esta en función, no solo de la pregunta formulada sino también de las opciones de respuesta que se dan. Por otra parte, se podrá comprobar aquí también, que algunos ítems figuran en más de un proceso. En este sentido, el esfuerzo por establecer distintos procesos de comprensión con el objeto de poder manejar una actividad tan compleja como la lectura, y proponer criterios e indicadores de evaluación diferenciados, no debe hacernos olvidar la interactividad del proceso en sí y la simultaneidad con la que estos procesos descritos se ponen en práctica. No obstante, para el análisis de datos, los ítems que figuran en más de una dimensión fueron adscritos finalmente a una de ellas, en función de la consideración de su mejor adecuación. De esta forma, para la prueba ECOS, los ítems 52, 54, 55 y 56 quedaron definitivamente encuadrados en la dimensión *Supervisión y autoevaluación del proceso de comprensión*. En el cuadro 28 mostramos las dimensiones contempladas en la prueba y la adscripción de cada uno de los ítems a los distintos procesos evaluados.

Validez y fiabilidad en la prueba ECOS

Aplicación piloto

La prueba de comprensión de textos ECOS también fue sometida a un proceso de aplicación piloto con el objetivo primordial de valorar diversos elementos relacionados con su estructura y su contenido que permitieran mejorar la prueba con vistas a garantizar al máximo su uso en el proceso de investigación. La aplicación piloto de la prueba ECOS para la valoración de la comprensión de textos en educación secundaria, al igual que ocurría con la escala de ajuste y con la prueba de comprensión elaborada para primaria, permitiría así mismo observar y analizar, previamente a su aplicación experimental con los sujetos de la muestra, diversos aspectos relativos a su validez y fiabilidad, así como al tiempo de ejecución estimado, la dificultad o facilidad de los ítems propuestos, la inteligibilidad de las instrucciones planteadas o el interés, la actitud y la motivación con que el alumnado cumplimentaba la prueba. Al igual que ocurría con la prueba de primaria, esta aplicación piloto se ha llevado a cabo en un solo centro educativo, el IES Dña Leonor de Guzmán, en Alcalá de Guadaíra (Sevilla). Los participantes formaban parte de dos grupos-clase de secundaria (43 sujetos), de 2º y 4º. Es importante destacar que estos grupos-clase han sido seleccionados de uno de los centros ya participantes en el estudio y en función de las posibilidades del investigador, de la orientadora del centro y de la disponibilidad que han mostrado los docentes. Dentro de estos grupos, lógicamente, también había alumnos y alumnas que presentaban dificultades de aprendizaje, algunos de los cuales también han participado de nuestro estudio. En este sentido, hemos optado igualmente por mantener a estos alumnos con su grupo a la hora de realizar esta prueba, con el objetivo de que los resultados obtenidos fueran representativos; si no en función del tamaño muestral, sí al menos en cuanto a la heterogeneidad habitual que caracteriza a la población escolar. En la tabla 24 mostramos una síntesis de la información estadística básica con relación a la aplicación piloto de la prueba ECOS.

El análisis exploratorio de los datos extraídos nos revela que, a diferencia de lo que ocurría con la prueba de comprensión elaborada para primaria, los índices de dificultad observados para la prueba ECOS muestran que la realización de ésta no ha representado un nivel de dificultad que pueda considerarse excesivo. Expresado como el porcentaje

Cuadro 28. Dimensiones contempladas en ECOS, procesos evaluados e ítems correspondientes

Dimensiones de la Comprensión	Procesos y elementos implicados en la lectura y que son evaluados con la prueba ECOS	Ítems
<i>Concepción sobre la lectura</i>	Para que sirve leer, en que consiste y que implica la lectura	37, 46
<i>Conocimiento de estrategias que pueden aplicarse</i>	Estrategias de planificación Estrategias de supervisión Estrategias de autoevaluación Estrategias de Transferencia y generalización Estrategias de integración textual Estrategias de integración texto-conocimientos	22, 31 4, 13, 14, 47 6, 36 36, 38, 47 25, 33 38
<i>Conocimientos previos</i>	Conocimientos sobre el tema del texto, sobre las situaciones planteadas y sobre los esquemas de acción descritos	10, 18, 28, 35, 40, 50, 53
<i>Integración textual</i>	Construcción de proposiciones e integración lineal de las ideas Distinción de la información relevante y construcción de proposiciones globales Jerarquización y secuenciación de la información Acceso a la información contenida en cuadros, tablas y esquemas Identificación y uso de la estructura textual	3, 15, 17, 27, 29, 30, 56 2, 19, 34, 52 7, 20 43, 44, 45 7, 20, 21, 24
<i>Integración del texto con los conocimientos y experiencias previas</i>	Establecimiento de relaciones Formulación de hipótesis y predicción Realización de inferencias Jucio de la intención del autor	32 8 1, 11, 12, 23, 32, 39, 41, 42, 49, 51, 54, 55 9, 16, 26, 48,
<i>Supervisión y autoevaluación del proceso de comprensión</i>	Detección de incoherencias y superación de las dificultades mediante la aplicación de estrategias de procesamiento léxico, sintáctico y semántico	5, 52, 54, 55, 56

promedio de alumnos y alumnas que equivocan sus respuestas, el dato indica que la prueba no ha resultado excesivamente difícil para el conjunto de participantes en esta aplicación piloto. Al igual que hicimos anteriormente con la interpretación de los resultados de la aplicación piloto de la escala de actitudes y de la prueba de comprensión de primaria, podríamos comprobar que este índice promedio de dificultad se incrementa de forma significativa cuando acometemos el análisis de los grupos con

Tabla 24. Indicadores estadísticos de la aplicación piloto de la prueba ECOS

Indicadores	Resultados de la aplicación piloto de ECOS
Muestra	43
Cursos	2º y 4º ESO
Edades	12 a 16 años
Ítems de la prueba ECOS	56
Índice de Dificultad medio	0,5125
Índice medio de Homogeneidad	0,3748
Puntuación media global	108,8908
Error típico de la media	2,1764
Desviación típica	14,2715
Mediana	104,1133
Varianza	203,676
Puntuación mínima	82,41
Puntuación máxima	135,94
Rango	53,53
Amplitud intercuartil	21,701
Asimetría	0,461
Curtosis	-0,663
Puntuaciones centiles	5 86,463 10 93,118 25 98,326 50 104,113 75 120,027 90 133,047 95 135,652
Prueba Kolmogorov-Smirnov de normalidad	0,157 (sig. 0,009)
Prueba Shapiro-Wilk (muestra < de 50 sujetos)	0,935 (sig. 0,027)
Fiabilidad (índice Alfa de Cronbach)	0,8906
Media chicas	106,0422 (DT = 10,8604; N = 18)
Media chicos	110,9418 (DT = 16,1993; N = 25)
Nivel educativo	
Media 2º ESO	102,8638 (DT = 8,9216; N = 22)
Media 4º ESO	115,2049 (DT = 16,1965; N = 21)

dificultades de aprendizaje en lectura. Las diferencias significativas con los resultados de aquellos alumnos sin dificultades de aprendizaje, nos indicarían la validez de la prueba en cuanto a discernir entre los resultados obtenidos por alumnos sin dificultades y los resultados del alumnado con dificultades, de los que, a priori, podrían esperarse menores puntuaciones. Hemos realizado un análisis mediante la prueba t para muestras independientes con el objetivo de contrastar las puntuaciones del alumnado de la muestra de aplicación piloto (AP) con las del conjunto de alumnos y alumnas con dificultades de aprendizaje que han formado parte, bien del GE o bien del GC1. En el análisis efectuado no se ha asumido la igualdad de varianzas, ya que los análisis previos habían revelado la no existencia de homocedasticidad. Los resultados muestran diferencias muy significativas a favor del grupo de sujetos que han participado de la aplicación piloto. Este aspecto confirma nuestra hipótesis y refuerza la validez de la prueba ECOS por cuanto el alumnado con dificultades en el aprendizaje de la lectura obtiene significativamente una menor puntuación en una medida del rendimiento en lectura. En la comparación que hemos realizamos, al igual que hiciéramos cuando realizábamos el contraste para la prueba de primaria, no se han eliminado de la muestra de AP a los alumnos con dificultades de aprendizaje que formaban parte de ella; aún así, también en secundaria las diferencias se muestran muy significativas a favor del conjunto de alumnos que realizó la aplicación piloto de la prueba.

	N	Media	DT	Prueba t	gl.	Sig.
Muestra de AP	43	108,8908	14,2715			
Muestra con DA	59	89,3517	9,0659	7,892	66,191	0,000
Total	102	97,5888	15,0304			

Por lo que respecta a los índices de homogeneidad de la prueba, hemos realizado los cálculos para cada ítem, obteniendo posteriormente el índice promedio. Los resultados obtenidos de estos cálculos han sido bastante positivos, mostrando un total de 25 correlaciones significativas al 0,01 y 12 al 0,05; 18 de las correlaciones son positivas aunque no llegan a ser significativas, y tan solo 1 presenta una correlación negativa, aunque sin llegar a ser significativa. Tanto el índice promedio de homogeneidad, como la mayoría de los índices para cada correlación, expresan que cada uno de los ítems de la prueba contribuye de manera significativa a evaluar lo mismo que la prueba en su conjunto. Por otra parte, para la prueba ECOS, al igual que ocurriera para su homóloga en primaria, hemos podido comprobar que las puntuaciones de la muestra de la aplicación piloto en su conjunto se adecuan a una distribución normal. En este caso, como puede apreciarse en la tabla, se ha obtenido el estadístico de la prueba de Shapiro-Wilk por tratarse de una muestra inferior a 50 sujetos. Por su parte, los índices de asimetría y curtosis expresan una distribución de las puntuaciones ligeramente asimétrica y con un valor negativo, indicando en este caso que hay un porcentaje importante de sujetos para los que la prueba no ha constituido una especial dificultad, obteniendo buenas puntuaciones en los resultados globales.

En cuanto al contraste efectuado entre las puntuaciones obtenidas por chicas y chicos en la prueba ECOS, podemos observar, a diferencia de lo que ocurría con los resultados del análisis de la escala APSL y de la prueba ECO 1, que la media obtenida por las chicas es inferior a la que obtienen los chicos, eso sí, sin que la diferencia observada llegue a ser estadísticamente significativa ($t = -1,187$; sig. 0,242, para 40,846 grados de libertad). Estos cálculos se realizaron sin asumir la igualdad de varianzas. Hemos realizado igualmente el cálculo de la prueba chi-cuadrado para cerciorarnos de que no existen

diferencias significativas entre el número de niñas y de niños que componen esta muestra de aplicación piloto; los resultados muestran la ausencia de diferencias relevantes ($\chi^2 = 1,140$; sig. 0,286, para 1 grado de libertad). En cuanto al nivel educativo, hemos confirmado la existencia de diferencias significativas tras realizar un contraste mediante la prueba *t* para muestras independientes y en ausencia de homocedasticidad ($t = -3,075$; sig. 0,004 para 30,802 grados de libertad). Los resultados son muy significativos a favor del grupo de alumnos y alumnas de cuarto de ESO.

De los análisis realizados previamente sobre los índices de dificultad y de homogeneidad en la prueba, hemos derivado otros más específicos que apuntan hacia un análisis más exhaustivo de cada uno de los ítems con el objetivo último de examinar la existencia de elementos que pudieran significar un obstáculo para una mayor validez interna y fiabilidad de la prueba en su conjunto. Este análisis lo venimos centrando en la excesiva dificultad o facilidad de algunos ítems, lo que, unido a la intensidad de la correlación con la puntuación global de la prueba, conduce a la valoración sobre la calidad de la información que aportan al conjunto de los resultados y, en su caso, a la consideración de su eliminación de la prueba. En la tabla 25 podemos observar los resultados de este análisis, que hemos organizado en función de las distintas dimensiones contempladas en la prueba. Por otra parte, la observación y el análisis de los datos expresados en la tabla nos revelan, tal y como apuntábamos anteriormente, que la prueba ECOS parece haber representado un nivel intermedio de dificultad para la mayor parte de los sujetos de la muestra piloto, con una gran mayoría de ítems (83,93%) cuyos índices promedio se encuentran entre el percentil 25 y el 75. Con respecto al resto de elementos, debemos destacar el ítem 18, que muestra un índice de dificultad muy bajo (4,65%). Podemos observar en la columna correspondiente a los índices de homogeneidad que, aunque sin llegar a ser significativa, este ítem mantiene una correlación positiva con la puntuación global bastante aceptable ($r = 0,245$). Cuando hemos procedido a simular la eliminación de este ítem, hemos podido comprobar que el índice de fiabilidad de la prueba no ha experimentado incremento con respecto al observado para el conjunto de los ítems contemplados inicialmente.

Estos análisis encaminados a comprobar el comportamiento de este índice de fiabilidad en el caso de eliminar de la prueba algunos de los ítems, nos muestran que, en general, los índices varían muy poco con respecto al nivel observado para el total, superando por muy poco el nivel citado en ocho ocasiones. La exploración de los ítems cuya hipotética supresión causaría un incremento del indicador de fiabilidad de la prueba, nos ha revelado que se trata de ítems de respuesta cuyo índice de homogeneidad no ha resultado estadísticamente significativo, encontrándose entre ellos el único que ha obtenido una correlación negativa con la puntuación total obtenida (ítem 6). En cuanto a los índices de dificultad de estos ítems, a diferencia de lo que ocurría con la prueba de primaria, en este caso no han resultado especialmente llamativos debido a que hayan supuesto una excesiva facilidad o dificultad para el alumnado; en este sentido, tan solo destacar el ítem 53, que ha resultado ser el que ha causado un mayor número de fallos entre el alumnado (88,37%). Por lo demás, el resto de estos ítems ha obtenido unos índices de dificultad que se encuentran dentro de los márgenes esperables en una prueba de este tipo, lo que, por otra parte, también puede ser aplicado al ítem 53. En cualquier caso, los índices de fiabilidad obtenidos tras la eliminación de estos ítems citados, no representaban un incremento que resultara significativo con respecto al obtenido con todos los ítems contemplados en la prueba. Este índice, como pudimos observar anteriormente ($\alpha = 0,8906$), está por encima de los niveles habitualmente aceptados y observados para pruebas semejantes a la nuestra; este hecho, unido a las ventajas que

Tabla 25. Índices de Dificultad (ID) e Índices de Homogeneidad (IH) para cada uno de los ítems de la prueba ECOS. (Los índices de dificultad están expresados en % de alumnos que fallan el ítem).

Dimensiones contempladas	Ítems	Índice de dificultad	Índice de homogeneidad
D1. Concepción sobre la lectura	37	74,42	0,272
	46	51,16	0,585**
D2. Conocimiento de estrategias que pueden aplicarse	4	60,47	0,136
	6	23,26	-0,028
	13	51,16	0,189
	14	74,42	0,343*
	22	76,74	0,260
	25	58,14	0,394**
	31	46,51	0,325*
	33	51,16	0,399**
	36	65,12	0,313*
	38	74,42	0,425**
D3. Conocimientos previos	47	58,14	0,612**
	10	62,79	0,475**
	18	04,65	0,245
	28	51,16	0,409**
	35	32,56	0,566**
	40	41,86	0,321*
	50	60,47	0,341*
53	88,37	0,115	
D4. Integración textual	2	44,19	0,618**
	3	41,86	0,336*
	7	32,56	0,454**
	15	25,58	0,499**
	17	34,88	0,363*
	19	55,81	0,318*
	20	60,47	0,194
	21	83,72	0,225
	24	72,09	0,236
	27	48,84	0,488**
	29	44,19	0,258
	30	37,21	0,266
	34	72,09	0,305*
	43	62,79	0,549**
44	48,84	0,651**	
45	44,19	0,498**	
D5. Integración del texto con los conocimientos y experiencias previas	1	20,93	0,256
	8	11,63	0,331*
	9	25,58	0,204
	11	18,60	0,432**
	12	32,56	0,403**
	16	18,60	0,340*
	23	46,51	0,464**
	26	39,53	0,259
	32	34,88	0,508**
	39	58,14	0,211
	41	69,77	0,489**
	42	51,16	0,647**
	48	51,16	0,647**
	49	74,42	0,703**
51	65,12	0,573**	
D6. Supervisión y autoevaluación del proceso de comprensión	5	44,19	0,297
	52	72,09	0,294
	54	62,79	0,159
	55	72,09	0,364*
	56	83,72	0,452**

(IH) ** La correlación es significativa al 0,01; * La correlación es significativa al 0,05

supone para la calidad de la información la diversidad de ítems en cuanto al grado de su dificultad, parecen no aconsejar la eliminación de ninguno de estos ítems. El análisis de los datos parecía indicar que la prueba en su conjunto otorgaba suficientes garantías con vistas a poder ofrecer información diferencial en función del alumnado participante y del momento de pasación de la prueba (pretest/postest); por otra parte, como hemos visto, los índices de homogeneidad en su conjunto también resultaron ser muy positivos, mostrando en general que la gran mayoría de los ítems contribuía de forma significativa a medir lo mismo que se está valorando con el conjunto de la prueba.

Hemos procedido igualmente a analizar las relaciones existentes entre los distintos procesos contemplados en la prueba. El comportamiento de las puntuaciones obtenidas en cada uno de los procesos y la posible interdependencia de estas entre sí o con respecto a la puntuación global obtenida, ha constituido un elemento de interés en función de la valoración de la estructura y el contenido de la prueba. Al realizar estos análisis por procesos, nuestro interés se ha centrado en conocer si las distintas dimensiones contempladas contribuían de forma conjunta a explicar los resultados y, por consiguiente, una mayor capacidad de comprensión lectora en los sujetos de la muestra. Por otra, parte, la constatación —en función de los datos obtenidos— de que la estructura de la prueba ayuda a interpretar más adecuadamente los resultados obtenidos, podía ser de gran valor a la hora de analizar las dificultades existentes en el alumnado de la muestra experimental, así como las posibles diferencias halladas entre los distintos grupos del estudio en función de su participación o no de las actividades del dispositivo de evaluación dinámica. Para ello, al igual que hicimos en anteriores ocasiones, efectuamos el análisis de una matriz de correlaciones entre las puntuaciones obtenidas en las distintas dimensiones de las que se compone la prueba ECOS. Hemos realizado igualmente los cálculos de la correlación con la puntuación total obtenida. Los resultados se muestran en la siguiente tabla.

DIMENSIONES CONTEMPLADAS EN LA PRUEBA ECOS

- D1 Concepción sobre la lectura.
- D2 Metaconocimiento de estrategias que pueden aplicarse.
- D3 Conocimientos previos
- D4 Integración textual.
- D5 Integración del texto con los conocimientos y experiencias previas.
- D6 Supervisión y autoevaluación del proceso de comprensión.

Tabla 26. Matriz de correlaciones entre las distintas dimensiones contempladas en la prueba ECOS

	D 1	D 2	D 3	D 4	D 5	D 6	Punt. total
D1	1,000	0,241	0,284	0,457**	0,566**	0,515**	0,600**
D2	0,241	1,000	0,371*	0,431**	0,525**	-0,070	0,643**
D3	0,284	0,371*	1,000	0,585**	0,614**	0,502**	0,751**
D4	0,457**	0,431**	0,585**	1,000	0,743**	0,324*	0,877**
D5	0,566**	0,525**	0,614**	0,743**	1,000	0,335*	0,910**
D6	0,515**	-0,070	0,502**	0,324*	0,335*	1,000	0,470**
Puntuación total	0,600**	0,643**	0,751**	0,877**	0,910**	0,470**	1,000

** La correlación es significativa al 0,01; * La correlación es significativa al 0,05.

La inspección de los datos de la matriz muestra que la gran mayoría de las correlaciones presentan significación estadística; concretamente, 15 correlaciones, entre las que se encuentran todas las relativas a la puntuación total, son significativas con un nivel de confianza del 99%; y 3 correlaciones presentan significación a un nivel del 95%. De las restantes, 2 correlaciones han resultado positivas, aunque no llegaron a ser significativas, y una presenta correlación negativa, aunque también sin significación estadística. El índice promedio de correlación entre las distintas dimensiones es de 0,4282. Por otra parte, la puntuación total obtenida en la prueba correlaciona muy significativamente con todas las dimensiones consideradas, mostrando unos niveles de correlación muy altos. Estos datos revelan una fuerte conexión entre los distintos procesos evaluados en la prueba, lo que a su vez informa sobre la unidad que presenta ésta con relación a los diversos bloques contemplados. En este sentido, la calidad de la comprensión de textos manifestada para cada una de estas dimensiones se reflejaría de forma significativa en la calidad de la comprensión en otras dimensiones, así como en la puntuación total, o, dicho de otro modo, la puntuación obtenida en cada una de las dimensiones explicaría en parte la puntuación en las otras. Sin embargo, es preciso hacer mención de la correlación negativa citada anteriormente. Ésta se ha producido entre la dimensión que evalúa el metaconocimiento estratégico y la dimensión que valora los procesos de supervisión y autoevaluación en la comprensión. La exploración del resto de correlaciones para estos dos bloques nos hace comprobar que su comportamiento con el resto de dimensiones ha sido positivo, manteniendo ambas correlaciones positivas y significativas en la práctica totalidad de casos, si bien, la dimensión *Supervisión y autoevaluación del proceso de comprensión* es la que mantiene una menor intensidad en la correlación obtenida con la puntuación total. A pesar de esto, el resultado es un tanto sorprendente; si recordamos, en la prueba elaborada para educación primaria, la correlación entre estos mismos bloques era muy significativa ($r=0,392$).

En cuanto a los ítems que mantenían una baja correlación con la puntuación total o que, en mayor o menor medida, habían resultado más fáciles o más difíciles que el resto, hemos comprobado que éstos se encuentran encuadrados principalmente en las dimensiones D2 (*Metaconocimiento estratégico*; ítems 4 ($r=0,136$), 6 ($r=-0,028$), y 13 ($r=0,189$)); D3 (*Conocimientos previos*; ítem 53 ($r=0,115$); y D6 (*Supervisión y autoevaluación del proceso de comprensión*; ítem 54 ($r=0,159$)). Podemos observar que precisamente en el bloque de metaconocimiento estratégico se encuentra el ítem que correlacionaba negativamente con el total de la prueba. Así mismo, constatamos que la mayor parte de estos ítems —aunque son pocos— están en las dimensiones que han correlacionado negativamente entre sí. Sin embargo, el análisis de la fuerza de las correlaciones de la puntuación total obtenida con cada una de estas dos dimensiones (0,643 con D2 y 0,470 con D6), así como las correlaciones igualmente significativas que mantienen con el resto de bloques contemplados, parece restar relevancia a la incidencia que haya podido tener la presencia de estos ítems en la versión aplicada de la prueba.

Al igual que hicimos con la escala APSL y la prueba ECO 1, hemos calculado el índice alfa de Cronbach para la prueba en función de sus dimensiones y hemos procedido a eliminar cada una de ellas para comprobar si la fiabilidad de la prueba experimentaba cambios cuando se prescindía de la influencia de uno de los bloques. Los resultados muestran que tan solo cuando eliminamos los bloques D2 y D6, es decir, *Metaconocimiento estratégico* y *Supervisión y autoevaluación del proceso de*

comprensión, el índice de fiabilidad experimenta un ligero incremento; por otra parte, el índice obtenido para el conjunto de dimensiones es bastante alto (0,8719), lo que unido a los datos anteriormente analizados y expresados en la matriz de correlaciones, así como al interés que para nosotros tenía precisamente la presencia de estas dos dimensiones, dirigidas a la evaluación del conocimiento sobre las propias estrategias de lectura y a la valoración de la puesta en práctica de estrategias de autorregulación, nos ha hecho mantener todos los bloques contemplados originalmente en la prueba. Tal y como hemos destacado anteriormente, los niveles de correlación hallados entre estas dimensiones y la puntuación global, apuntan a su relevancia final en el resultado obtenido por los sujetos en la prueba, reforzando así mismo las relaciones entre el contenido evaluado en estas dimensiones y la comprensión efectiva de los textos. Por otra parte, nos interesaba especialmente comprobar el comportamiento de la muestra experimental, compuesta por alumnado con dificultades de aprendizaje, en estos bloques y con relación al resto de procesos. Nuestra hipótesis en este sentido se dirigía hacia menores puntuaciones en ambas dimensiones, al tiempo que una menor correlación con la puntuación global, quedando por ver el comportamiento de la correlación entre éstas.

Al igual que hiciéramos con las puntuaciones de la escala de actitudes y con las de la prueba de comprensión de textos para primaria, las puntuaciones de la prueba ECOS también han sufrido una transformación similar. Las razones que nos han llevado a realizar esta conversión de las puntuaciones directas en típicas, así como la posterior transformación en una escala derivada con media en 100 y desviación típica en 16, son las mismas que ya expusimos para las dos pruebas ya descritas y analizadas con anterioridad.

La validación de expertos

Para la prueba ECOS procedimos igualmente a su validación por parte de los expertos. En este caso, el análisis pudo ser realizado por un solo profesional del ámbito educativo, una orientadora que desempeñaba su trabajo en un IES. Al igual que ocurriera con el grupo de expertos que valorara la prueba de comprensión para educación primaria, esta profesional había participado ya del análisis de la escala APSL. A pesar de que en un principio pensábamos contar con un pequeño grupo de expertos también para la prueba en educación secundaria, finalmente esto no fue posible, por lo que la selección de esta profesional se realizó en virtud de sus cualidades con relación al tema objeto de valoración, así como en función de su disponibilidad e implicación. En cuanto al calendario aproximado para esta validación, éste no difiere en absoluto de los que ya hemos expuesto para el resto de las pruebas, constando de los siguientes momentos: (a) elaboración del material que sería distribuido entre los colaboradores para realizar la valoración; (b) selección de expertos; (c) proceso de validación; (d) recogida y análisis de los datos; (e) conclusiones y modificaciones en función del proceso de validación; y (f) versión definitiva de la prueba.

En cuanto al material utilizado para el proceso de validación (anexo VII), éste ha estado compuesto por:

- Una plantilla que recogía la totalidad de los ítems de la prueba, clasificados a su vez en función de las distintas dimensiones contempladas, así como una escala de valoración cuantitativa donde los expertos debían marcar de 1 a 4 en función

de la adecuación de cada uno de los ítems a la evaluación de las distintas dimensiones contempladas en la prueba. En la plantilla se recogía así mismo una gradación cualitativa que permitía comprender y realizar la valoración, así como interpretar posteriormente las marcaciones efectuadas. Esta gradación contenía cuatro niveles, a los que ya hicimos referencia cuando analizábamos la escala de actitudes y la prueba de comprensión de primaria:

(1) *Nivel bajo o muy bajo*. El nivel de adecuación del ítem para la evaluación de los procesos descritos es bajo o muy bajo.

(2) *Nivel medio-bajo*. El nivel de adecuación no es satisfactorio. El contenido del ítem se muestra claramente insuficiente para evaluar los procesos contemplados.

(3) *Nivel medio-alto*. El nivel de adecuación del ítem para la evaluación de los procesos descritos es satisfactorio. El contenido del ítem evalúa dichos procesos.

(4) *Nivel alto o muy alto*. El nivel de adecuación es notable. El contenido del ítem evalúa de forma adecuada o muy adecuada los procesos descritos.

Esta plantilla también recogía una escala de valoración cuantitativa donde los expertos (en este caso uno solo) debían valorar globalmente la prueba ECOS, otorgando así mismo de 1 a 4 puntos en función de su adecuación para evaluar la comprensión de textos en educación secundaria.

- Un cuestionario de preguntas abiertas sobre diversos aspectos de la prueba, donde la experta debía realizar una valoración cualitativa acerca de estos aspectos y donde podía aportar todas las sugerencias que considerara oportunas sobre los elementos integrantes de la prueba en su conjunto para la mejora del instrumento y para su adecuación a los objetivos del estudio. Esta valoración cualitativa servía al propósito de obtener más elementos de validez externa del contenido y de la adecuación global de la prueba a los objetivos de la investigación. Las cuestiones sobre las que debía realizarse esta valoración cualitativa fueron estimadas especialmente relevantes para la consideración de la bondad de la prueba. La experta era informada a través de las instrucciones recogidas en la plantilla sobre la que debía realizar su valoración, de que podría expresar ésta en forma de dudas, comentarios críticos sobre aspectos que deberían ser mejorados, valoraciones positivas, valoraciones sobre otros aspectos no recogidos en la plantilla y que se consideraran relevantes, etc. Los elementos propuestos para su valoración eran los siguientes:

* Los textos (adecuación y contenido, expresión, tipologías de textos y estructuras contempladas, vocabulario...). Observaciones, dudas, comentarios.

* Las dimensiones de la comprensión contempladas.

* La prueba contemplada en su totalidad.

El método de análisis utilizado en la valoración que hemos efectuado del trabajo desarrollado por la experta participante, ha sido semejante al que expusimos con anterioridad en nuestro análisis de la escala de actitudes y de la prueba de comprensión para primaria. Tal y como apuntábamos entonces, el aspecto más relevante de este análisis es su carácter esencialmente cualitativo y su orientación a la mejora del instrumento analizado. Posteriormente recogemos algunos de los datos y conclusiones

que podemos extraer de la valoración realizada, en este caso, por esta única profesional. Por otra parte, con relación a los datos provenientes del análisis más cuantitativo, en la siguiente tabla podemos observar una síntesis de las valoraciones efectuadas. La puntuación expresada en la columna de la experta responde al promedio efectuado entre sus valoraciones de la adecuación de cada uno de los ítems contemplados en la prueba.

Tabla 27. Valoraciones cuantitativas del contenido de los ítems de la prueba ECOS en función de los procesos contemplados (E = Experta)

Dimensiones de la comprensión contempladas	E 1 y Total
D1. Concepción sobre la lectura	4,00
D2. Metaconocimiento de estrategias que pueden aplicarse	2,57
D3. Conocimientos previos	3,14
D4. Integración textual	3,85
D5. Integración del texto con los conocimientos y experiencias previas	4,00
D6. Supervisión y autoevaluación del proceso de comprensión	3,80
Promedio de las valoraciones	3,56
Valoración global de la prueba ECOS	4

La inspección de los datos muestra que, en general, la valoración otorgada por esta profesional al conjunto de los ítems de la prueba acerca de su adecuación para evaluar los distintos aspectos de la comprensión, en función del proceso en el que se hallaban incluidos, ha sido muy positiva. Como podemos observar, el promedio de sus valoraciones supera ampliamente el nivel medio-alto y su valoración global acerca de la prueba ECOS en su conjunto, ha sido la máxima posible. Teniendo en cuenta estos datos, aunque sin olvidar que se trata de una sola valoración, podemos concluir que el nivel general de adecuación de los ítems para evaluar los procesos contemplados y considerados en la prueba ECOS, ha sido considerado más que satisfactorio y que, por tanto, el contenido de estos ítems evalúa estos procesos. Sin embargo, esta valoración positiva global no se ha dejado sentir por igual en todas las dimensiones evaluadas. En el bloque de metaconocimiento estratégico, podemos observar que la valoración baja ostensiblemente con relación al resto. En la plantilla trabajada por la profesional que realizó la valoración, hemos podido contar hasta cinco ítems para los que esta experta no emitía un juicio cuantitativo, anotando diversos comentarios que valoraban cualitativamente la adecuación del ítem. Tal y como hemos realizado anteriormente con las valoraciones correspondientes a las pruebas ya analizadas aquí, hemos otorgado 1 punto, es decir, la valoración más baja posible, a cada ítem no valorado, con objeto de poder realizar el promedio. Este hecho ha ocasionado que el promedio calculado haya sido especialmente bajo; por otra parte, la mayoría del resto de valoraciones de la experta para los ítems de este bloque se caracteriza por puntuaciones de nivel medio-alto o alto.

El análisis de los datos también nos revela que, a excepción de la ya comentada puntuación para la dimensión 2, el resto de valoraciones se encuentra por encima del nivel medio-alto de puntuación (3) y el 66,67% supera la puntuación promedio de 3,50. Destacar así mismo que el ítem 6 (aquel que presentaba correlación negativa con el total de la prueba) ha sido valorado con la máxima puntuación por esta profesional, precisamente dentro del bloque de ítems que evalúan metaconocimiento estratégico. También el ítem 53, que ha representado la máxima dificultad para el conjunto de sujetos de la muestra piloto, ha obtenido la máxima puntuación en esta valoración de

expertos. Parece ser que, a priori, estos ítems han sido positivamente valorados en función de su adecuación para evaluar los conocimientos sobre las propias estrategias de comprensión, así como los conocimientos previos sobre el contenido de los textos trabajados. Sin embargo, la puesta en práctica de la prueba mediante la aplicación piloto, como hemos podido comprobar, nos mostró la ausencia de conexión del primero de los ítems con el total de la puntuación obtenida, así como el incremento —aunque leve— del índice de fiabilidad cuando eliminábamos el ítem de la prueba; este leve incremento de la fiabilidad también se daba cuando eliminábamos del conjunto al ítem 53, que presentaba una excesiva dificultad con relación al resto de los elementos de la prueba. Algunos ítems, al igual que hiciéramos en la prueba de primaria, fueron encuadrados originalmente en varias dimensiones a la hora de la valoración de expertos, por estar relacionados con todas ellas, tal como se muestra en el cuadro 28; ha sido posteriormente, tal y como comentábamos en páginas anteriores, en función de las revisiones efectuadas tras la aplicación piloto y la validación de expertos, y ante la necesidad de encuadrarlos definitivamente en algún bloque a efectos del análisis de datos, cuando han sido clasificados en un determinado proceso.

La valoración cualitativa realizada por la profesional que ha participado de esta valoración de expertos, a través de las preguntas abiertas del cuestionario, ha dejado entrever aspectos relevantes que han sido objeto de análisis por nuestra parte con el doble objetivo de mejorar en lo posible la prueba ECOS antes de su aplicación en la fase experimental del estudio, o, al menos, de conocer posibles carencias o deficiencias en la misma que pudieran ser de esta forma tenidas en cuenta y controladas en los análisis posteriores. Ante las valoraciones efectuadas, hemos realizado un análisis del contenido de cada una de las respuestas y valoraciones realizadas en función de las cuestiones planteadas. Al igual que hicimos cuando analizamos la escala de actitudes y la prueba de primaria, exponemos las observaciones, dudas o comentarios críticos de la única profesional para, posteriormente, centrarnos en aquellos aspectos que se vieron afectados por estas valoraciones en la versión final de la prueba.

* Sobre los textos (adecuación y contenido, expresión, tipologías de textos y estructuras contempladas, vocabulario...). Observaciones, dudas, comentarios.

Se hace alusión al hecho de que pensando en sujetos con dificultades de aprendizaje de la lectura, los textos pueden resultar complejos. La orientadora recoge que su experiencia piloto previa a la valoración cualitativa de la prueba, le hizo constatar que los alumnos y alumnas participantes decían encontrar más dificultad en el texto número 8 y en la actividad planteada con el texto número 4. El análisis de los índices de dificultad de la muestra piloto confirma solo en parte la valoración de la orientadora; en efecto, para el texto 4, de los cuatro ítems que siguen al texto, tres presentan índices de dificultad altos (ítem 21: 83,71; ítem 22: 76,74; e ítem 24: 72,09), mientras que el ítem 23 presenta un índice de dificultad de 46,51. Por su parte, en el texto 8, los índices de dificultad observados para los tres ítems que siguen al texto no han sido altos (ítem 43: 62,79; ítem 44: 48,84; e ítem 45: 44,19).

* Sobre las dimensiones de la comprensión contempladas

En este caso, el análisis de la experta solo hace referencia a que las dimensiones contempladas en la prueba son las adecuadas.

* Sobre la prueba contemplada en su totalidad

Se expresa en esta valoración una preocupación acerca del nivel de la prueba elaborada con relación a la muestra de alumnos y alumnas que conformarán los grupos de la investigación. En este sentido, se expresa que el nivel de comprensión de algunos alumnos de primer ciclo de secundaria y de los alumnos con dificultades de aprendizaje, estaría más próximo a los textos y las actividades planteadas en la prueba de primaria (la orientadora conocía esta prueba). Indicaba esta profesional que podría darse el caso de que por la dificultad de estos alumnos para acceder a la comprensión de los textos de la prueba ECOS, ésta no se mostrara sensible a los cambios derivados de la intervención con el dispositivo de evaluación dinámica.

Por otra parte, también se alude a una posible desproporción en el número de ítems que conforman cada uno de los procesos evaluados. En este sentido se apunta a la posibilidad de haber reducido el número de ítems que valoran algunos procesos (*Conocimientos previos*, *Integración textual* —construcción de proposiciones e integración lineal de las ideas— e *Integración del texto con los conocimientos y experiencias previas* —realización de inferencias—). En línea con estos comentarios acerca del número de ítems, hacemos referencia a otros que la misma profesional realizó en la hoja de valoración cuantitativa de los ítems; en estos comentarios, anotados junto a los ítems valorados, la orientadora aludía sobre todo, en forma de duda o de observación, a la posibilidad de que el ítem en cuestión pudiera encuadrarse mejor en otro de los procesos contemplados.

Por último, la valoración cualitativa contenía una propuesta relacionada con la plantilla de corrección de la prueba. Se recogía una sugerencia para facilitar la identificación gráfica de aquellos procesos afectados en un sujeto una vez que se procedía a la corrección de la prueba. Para conseguir esto, se proponía realizar una agrupación de las preguntas por bloques para que el análisis fuera más fácil y se accediera, no solo a una interpretación cuantitativa de los aciertos y errores cometidos por el alumno, sino también a una interpretación cualitativa en función de los procesos en los que fallaba.

Repercusiones de la validación de expertos para la versión final de la prueba ECOS

Las repercusiones de la valoración efectuada para la construcción de la prueba ECOS con vistas a su utilización en la fase experimental del estudio, han sido notables. Como apuntábamos anteriormente, parte de estas repercusiones no han tenido que evidenciarse necesariamente en cambios concretos que afectaran a la estructura o al contenido de la prueba de forma previa a su aplicación; más bien, en este caso, se ha tratado de tomar en consideración las excelentes aportaciones y observaciones de la profesional que ha participado en la valoración con el objetivo de tenerlas en cuenta en los análisis que posteriormente se han realizado entre los grupos del estudio en las distintas fases del estudio. En efecto, debido a algunos de los comentarios críticos de esta orientadora, se ha procedido a analizar los resultados tomando en cuenta la posibilidad de un efecto techo para el alumnado con dificultades de aprendizaje, o la posibilidad de que algunos textos hubieran representado una especial dificultad para el conjunto de sujetos participantes.

Al igual que realizamos con las pruebas anteriores, la valoración efectuada por esta experta fue analizada y discutida con ella de forma conjunta. En este análisis, valorado

también con la posterior aplicación piloto de la escala, se llegó a reforzar algunas de las decisiones relativas a la estructura de la prueba, así como a su contenido. Por otra parte, con respecto a los comentarios relativos a la excesiva dificultad de la prueba con vistas a su utilización para mostrar las diferencias entre los grupos experimental y control en función de su participación del tratamiento, pudimos comprobar a través de los datos ofrecidos por la aplicación piloto, que los índices de dificultad de la prueba no habían resultado excesivamente altos, lo que no hacía temer especialmente que la prueba no se mostrara sensible a la hora de captar los posibles cambios derivados de la intervención con el dispositivo EDPL. En cambio, estos análisis críticos nos fueron muy útiles para analizar los resultados de las muestras piloto y experimental en la prueba de comprensión elaborada para primaria, ya que la posibilidad apuntada por la orientadora cobraba un mayor sentido a tenor de los índices de dificultad obtenidos para la muestra piloto con ECO 1.

Por otra parte, nos es válido el comentario que realizamos cuando analizábamos las valoraciones efectuadas acerca de la prueba de primaria, en el sentido de que si bien muchos de los sujetos con dificultades de aprendizaje podrían tener problemas a la hora de finalizar la prueba debido a su longitud o complejidad, la estructura de la misma, con diferentes textos y cuestiones para cada uno de los textos, así como la distribución de los ítems de respuesta correspondientes a las distintas dimensiones contempladas a lo largo de la prueba, hacían que el objetivo de la aplicación de la prueba pudiera verse justificado y cumplido. La extensión y complejidad de la prueba había sido fruto de dos consideraciones fundamentales, de una parte, la necesidad y conveniencia de incluir todos los procesos implicados en la lectura que pudieran ser evaluados en un sistema de prueba cerrada, y de otra, la posibilidad que ofrecía de mostrar información acerca de un amplio abanico de niveles de competencia del alumnado que participaría en la investigación. Aunque éramos conscientes de la dificultad que podría plantear la prueba, especialmente para el alumnado con dificultades de aprendizaje, optamos en este caso por no modificar su contenido ni reducir el número de textos o ítems. La prueba, al igual que la elaborada para primaria, había sido ya modificada previamente a su validación por parte de la profesional que realizó la valoración de expertos, y las acertadas y relevantes observaciones efectuadas por ésta habían sido ya objeto de discusión previa. La decisión parecía acertada en función de la posibilidad que ofrecía una prueba extensa y que contemplara diferentes tipologías textuales y procesos a evaluar, de captar información relevante tomando en consideración la diversidad de capacidades relacionadas con la comprensión, así como de establecer diferencias entre la ejecución del alumnado con problemas de comprensión en el marco de unas referencias estándar ofrecidas por la ejecución de la muestra piloto.

Otros instrumentos: los cuestionarios EIDAL y EIDAL-MP sobre expectativas e ideas acerca de las dificultades de aprendizaje de la lectura

A lo largo del desarrollo de la investigación, se han aplicado también otros instrumentos que nos han permitido ampliar el conocimiento sobre las circunstancias y condiciones que rodearon la intervención educativa en las dificultades lectoras de los alumnos y las alumnas participantes en el estudio. Su aplicación nos ha acercado a las expectativas y el pensamiento del profesorado y de las familias del alumnado participante acerca de la educación y de la acción educativa emprendida, o por emprender, con relación al alumnado con dificultades de aprendizaje en la lectura. Estas pruebas se han pasado una sola vez, durante la fase experimental del estudio. Se trata de

dos cuestionarios elaborados especialmente para esta investigación: un cuestionario para el profesorado de los niños y las niñas que participan en el estudio, sobre las expectativas e ideas acerca de la evaluación y la intervención educativa en las dificultades de aprendizaje de la lectura (EIDAL), y otro pensado para las madres y padres del alumnado, para valorar también las ideas y expectativas que tienen con relación a las dificultades lectoras de sus hijos, (EIDAL-MP). Hemos incluido la pasación de estas dos pruebas en el grupo control, aunque no exista la posibilidad de establecer conexiones entre los resultados de las mismas y los de la aplicación del proceso de evaluación dinámica. A parte de otorgarnos la posibilidad de establecer un cierto control de la variable: ideas y expectativas de los profesores y padres con respecto a las dificultades de aprendizaje de la lectura, control establecido entre profesorado, o familia, de alumnos que están encuadrados en el grupo experimental y profesorado o familia de alumnos que pertenecen al grupo control; nos ha permitido también ampliar la muestra de profesores y padres que explicitan ideas y expectativas sobre las dificultades de aprendizaje de la lectura, lo que, sin duda, puede contribuir a enriquecer nuestro estudio.

Se trata de cuestionarios de preguntas abiertas, en las que el profesorado o la familia puede expresar sus ideas, concepciones, expectativas y procedimientos habituales con relación a la atención educativa del alumnado con dificultades en la lectura. Precisamente por tratarse de preguntas abiertas, la cumplimentación quizás haya conllevado el empleo de un mayor tiempo, lo que puede haber dificultado un poco la necesaria colaboración de todo el profesorado, así como de algunas familias. Por otra parte, la elección de este modelo de cuestiones en detrimento de la presentación de preguntas con opciones cerradas de respuesta, respondía a la idea de obtener una mayor riqueza en las respuestas, otorgando mayores posibilidades de analizar cualitativamente la información obtenida. Su valoración sería, por tanto, sobre todo cualitativa y nos serviría, entre otras cosas, para poder establecer comparaciones y relaciones entre aspectos como: (a) acciones educativas emprendidas desde la escuela para mejorar el rendimiento en lectura del alumnado con dificultades de aprendizaje de lectura y rendimiento efectivo de estos; (b) ideas y expectativas del profesorado acerca de las posibilidades de aprendizaje y avance de estos alumnos, y el desempeño de éstos en las actividades escolares relacionadas con la lectura; (c) ideas y expectativas de los padres/madres acerca de la educación en general, y de la lectura en particular, y rendimiento, motivación y conducta del alumnado; etc. Su valoración también nos serviría potencialmente para comparar sus resultados con aquellos obtenidos en la evaluación dinámica de los procesos implicados en la lectura, ya que, a lo largo de su puesta en práctica, se habrían valorado aspectos relacionados con los procesos de enseñanza y aprendizaje llevados a cabo en el aula donde habitualmente se encuentra el alumno estudiado y, por otra parte, también se valorarían aspectos relacionados con el contexto social y familiar en que se desenvuelve el alumno.

Estructura y contenido de los cuestionarios EIDAL y EIDAL-MP

El cuestionario EIDAL (Anexo XI) se compone de 26 preguntas abiertas que versan sobre diversos aspectos relacionados con la evaluación y la acción educativa con el alumnado que presenta dificultades de aprendizaje en la lectura. Aunque las preguntas se presentan de forma más o menos aleatoria, hemos procurado situar de forma precisa determinadas preguntas en función de que pudieran introducir algunas cuestiones que se tratarían, o a las que se aludiría, posteriormente. El cuestionario se estructura en siete

bloques en función del contenido de las cuestiones que plantea. Esta estructura junto con los ítems correspondientes se presenta en la siguiente tabla.

Tabla 28. Estructura e ítems en el cuestionario EIDAL

Dimensiones contempladas en EIDAL	Cuestiones
A qué se deben las dificultades de aprendizaje. Factores que influyen	1, 6, 9, 12, 13
Metodología en la enseñanza/facilitación y evaluación de la lectura/comprensión	2, 4, 8, 15, 17, 19, 21
Aspectos/elementos/procesos implicados en la comprensión y en su mejora	3, 18, 22, 26
Colaboración con otros profesionales y con las familias	5, 20
La lectura en el currículum. Organización de la enseñanza de la lectura/comprensión. Cuándo debe ser trabajada la lectura	7, 23
Expectativas, valoración e ideas sobre el aprendizaje de alumnos con NEE	10, 11, 14, 16
Conocimiento y trabajo con programas de mejora de la comprensión de textos	24, 25

Por su parte, el cuestionario EIDAL-MP (Anexo XII) está compuesto por 23 preguntas abiertas, dirigidas a las familias del alumnado participante en el estudio, que abordan contenidos relativos a las expectativas e ideas relacionadas con la lectura y las dificultades de su aprendizaje. Al igual que comentábamos anteriormente, las cuestiones han sido planteadas de forma aleatoria, aunque cuidando la colocación de determinadas preguntas que introducen un tema que será abordado posteriormente en otra cuestión. La estructura del cuestionario, así como los ítems que se corresponden con los distintos bloques contemplados, está expuesta en la siguiente tabla.

Tabla 29. Estructura e ítems en el cuestionario EIDAL-MP

Dimensiones contempladas en EIDAL-MP	Cuestiones
Actuación ante diversos supuestos prácticos relacionados con las dificultades de aprendizaje de la lectura	1, 2, 3, 5, 7, 11, 19, 21
Ideas acerca de la implicación que debe tener la familia y posibilidades de intervención o de influencia para la mejora de las dificultades de aprendizaje	4, 15
Expectativas y deseos acerca de los estudios o la labor que realizará su hijo/a	6, 18
Información sobre disposición de elementos materiales (libros infantiles/juveniles)	8
Experiencia/contacto de la madre/padre con la lectura actualmente	9
Ideas sobre participación conjunta en actividades de lectura o audiovisuales	10, 20
Factores que influyen en las dificultades de aprendizaje	12, 23
Concepción de la lectura y de su relevancia	13, 14, 16
Expectativas sobre el aprendizaje de alumnos que muestren dificultades de aprendizaje o NEE	17
Colaboración con el profesorado	22

6.2 El dispositivo de evaluación dinámica de procesos implicados en la lectura (EDPL)

Lo que es y lo que no es el dispositivo EDPL

El dispositivo de *Evaluación Dinámica de Procesos implicados en la Lectura* (EDPL) podría definirse como una estrategia de evaluación dinámica y continua de los procesos implicados en la actividad lectora, con especial atención a los procesos metacognitivos, ya que son este tipo de procesos los que a nuestro juicio parecen tener un mayor peso en la explicación de las dificultades de aprendizaje de la lectura, y de forma más concreta en lo relativo a los procesos de comprensión. Este dispositivo no es un test convencional de lectura, tal y como se han descrito estas pruebas en el capítulo uno de este trabajo; tampoco podría considerarse una prueba estandarizada que admite una valoración cualitativa; por otra parte, no es un procedimiento de ED que responda a un modelo *pretest/ instrucción /postest* como el LPAD de Feuerstein o la Evaluación del Potencial de Aprendizaje de Budoff o Fernández-Ballesteros. Este dispositivo no encaja en este modelo ni tampoco en el de *pretest/ tarea-retroalimentación continua /postest* de Campione y Brown o en el de condiciones de prueba o *prueba de los límites* propuesto por Carlson y Wiedl. Se trata básicamente de un modelo de ED de los procesos de lectura y comprensión en el que se pretende recabar información útil para la intervención y, por consiguiente, para la mejora de los problemas de aprendizaje planteados. La intervención del evaluador tiene lugar en el transcurso del proceso y puede llegar a establecer ayudas hasta conseguir que el alumno comprenda y realice una determinada tarea. Uno de los objetivos no explícitos de los procesos de ED y, sin embargo, una de sus consecuencias más relevantes, consiste en proporcionar una conexión entre evaluación e intervención. El modo óptimo de conocer la capacidad de respuesta de un sujeto a la intervención es evaluar los efectos de cambios inducidos experimentalmente sobre su actuación en procesos estrechamente relacionados con la instrucción (Swanson, 2000). Esta conexión se facilita además por características como: (a) un papel activo del evaluador; (b) una interacción colaborativa entre el evaluador y el sujeto evaluado; y (c) un esfuerzo deliberado para introducir cambios en los procesos evaluados con el fin de obtener información relevante sobre el tipo de ayuda que el alumno necesita para resolver con éxito la actividad. Esta conexión entre los procesos de evaluación e intervención, así como los elementos que la hacen posible, caracterizan la propuesta de evaluación dinámica recogida en el dispositivo EDPL.

Como hemos mencionado anteriormente, el dispositivo EDPL no tiene una estructura *pretest/ mediación /postest*; sin embargo, este formato ha sido utilizado de forma general para el diseño de la investigación presentada, debido a que permitía diferentes tipos de análisis considerados relevantes en función de los objetivos del estudio. El dispositivo podría admitir este tipo de formato; de hecho, en algunas de las actividades propuestas, se han analizado resultados en función de este patrón. Por otra parte, lo relevante de nuestra propuesta para el contexto escolar consiste en evaluar el *proceso de aprendizaje* y esto, tal y como hemos apuntado ya con anterioridad, no se consigue satisfactoriamente con los resultados obtenidos en un *postest*, aunque éste forme parte de una prueba dinámica; ni estableciendo exclusivamente las diferencias entre las actuaciones *pretest* y *postest* de una prueba estática.

El dispositivo EDPL se aplica, no solo con el fin de obtener comparaciones con pruebas estandarizadas de rendimiento en lectura; tampoco se pone en práctica con la única

intención de comprobar si su aplicación, junto con la intervención propuesta, conlleva posteriores cambios en el sentido de ganancias en rendimiento lector, medido con las pruebas estandarizadas. Además de todo lo anterior, el EDPL se aplica para comprobar y analizar su eficacia y utilidad como propuesta de evaluación e intervención ante las dificultades de aprendizaje de la lectura. En este sentido, si ayuda a los orientadores, profesores de apoyo, logopedas, tutores, etc. en el diagnóstico de las dificultades y de dónde se encuentran éstas, a encontrar, mediante el análisis del proceso de resolución de cada una de las actividades propuestas, el modo en el que mejor podemos apoyar al alumno para mejorar su lectura; entonces esta estrategia de evaluación-intervención será válida. Después habrá que analizar si es costosa en tiempo y energía —aspectos nada desdeñables, por otra parte— pero serán éstos aspectos que podrán mejorarse posteriormente. Lo relevante, a nuestro juicio, es comprobar si una puesta en práctica lo más adecuada posible a las propuestas metodológicas que realizamos, ocasiona la satisfacción en los aplicadores debido a la constatación de haber encontrado un conjunto de instrumentos y de patrones metodológicos de actuación, que contribuyen de manera efectiva y eficaz a detectar *dónde* se encuentran las dificultades, al tiempo que *orientan la actividad hacia la mejora de esas dificultades*.

Con la aplicación del dispositivo EDPL no pretendemos cuantificar la distancia entre lo que el sujeto es capaz de realizar de forma autónoma y aquello que es capaz de hacer con la guía y la colaboración de un adulto o compañero aventajado; es decir, nuestra intención no ha sido la de cuantificar la ZDP del alumno evaluado. El objetivo es, más bien, *actuar en la ZDP* para: (a) obtener información sobre lo que el alumno es capaz de realizar con la guía y la mediación del evaluador, es decir, información sobre sus procesos de comprensión en desarrollo; (b) obtener información sobre el tipo y el grado de ayuda que le hace falta para comprender y ejecutar la tarea propuesta; (c) evaluar el proceso de aprendizaje, ya que la ED contempla por principio la inclusión de elementos propios de la intervención que constituyen la posibilidad de observar procesos de establecimiento de relaciones y asociación, incorporación y construcción de aprendizajes; (d) determinar qué tipo y grado de apoyo es idóneo para facilitar y optimizar el aprendizaje; y (e) determinar dónde deben situarse los próximos objetivos de la intervención educativa en función de los límites actuales de su ZDP.

Construcción del dispositivo y aplicación piloto

El trabajo de investigación realizado durante las sesiones pertenecientes a un Seminario de post-grado sobre Comprensión Lectora, llevado a cabo durante los cursos 97/98 y 98/99 en el Departamento de Psicología Evolutiva y de la Educación de la Universidad de Sevilla, y bajo la coordinación del profesor Dr. Joaquín Mora Roche, dio como resultado un material elaborado tras la exposición de cada uno de los grupos de trabajo formados y que recogía aquellas aportaciones, bien elaboradas por los propios grupos o bien recogidas de otros autores, consideradas más relevantes para la evaluación y la intervención en dificultades asociadas a los procesos de lectura. Este material fue la base inicial sobre la que fuimos desarrollando un conjunto de actividades y de patrones metodológicos de actuación que respondieran de una manera lo más completa posible a una evaluación dinámica del proceso lector desde una perspectiva metacognitiva (Navarro, Lama y Molina, 2001). En primer lugar, es necesario decir que el tema central del citado Seminario fue la metacognición y, más concretamente, las habilidades metacognitivas implicadas en la comprensión lectora. Por tanto, el tener en cuenta este tipo de habilidades, el buscar su implicación en los procesos de lectura, así como la

mejor forma de evaluarlas y de contribuir a su optimización, ha sido nuestra tarea primordial durante este tiempo. El nacimiento y desarrollo de este trabajo tuvo lugar durante sus estadios iniciales en un ámbito de formación de profesionales de la Psicología y la Educación, concretamente en el ámbito de la Universidad, en un contexto de trabajo de pequeño grupo, donde la facilitación y el fomento de la discusión, la creatividad, los procesos de investigación-acción, la valoración crítica y la exposición de las propias ideas, se constituían en características esenciales. Precisamente estos elementos son, a nuestro entender, básicos para favorecer el desarrollo de cualquier proceso de aprendizaje en el alumnado y también lo son para favorecer el desarrollo de propuestas que, como la que se realiza desde EDPL, tratan de contribuir a la mejora de los procesos de evaluación e intervención relacionados con la lectura.

El trabajo desarrollado durante cuatro años por nuestro pequeño grupo de investigación ha contemplado una revisión continua del material que se iba elaborando. Esta revisión respondía a nuestro interés por adecuar las propuestas teóricas, expuestas o realizadas durante las sesiones del Seminario, a aquello que queríamos evaluar, y también por adecuar las actividades seleccionadas o desarrolladas, a estas propuestas teóricas. Esto ha conllevado que se modificara sustancialmente durante estos años nuestra propuesta inicial. De esta forma se han ido generando distintas actividades de evaluación-intervención, sometidas a la revisión continua de cada uno de los componentes del grupo —y puestas en práctica, en la medida de las posibilidades, en nuestros propios centros de trabajo—, que han tratado de responder a cada una de las dificultades posibles en cada uno de los procesos estudiados.

El trabajo se estructura básicamente en dos partes. En una primera parte, que consta de seis capítulos, se aborda la introducción y encuadre de su contenido, exponiendo las bases teóricas que sustentan el trabajo y desarrollando los elementos básicos y las premisas que constituyen la esencia de la propuesta: *una opción clara por evaluar procesos; evaluación dinámica del proceso de resolución de la actividad de lectura; análisis preferente de procesos metacognitivos; evaluación del contexto de aula; y contextualización de los instrumentos de evaluación*. En la segunda parte del trabajo, que contiene tres capítulos, se expone nuestra propuesta de Evaluación Dinámica de procesos en la lectura. Para ello, en el capítulo siete se detallan los procesos implicados que se han considerado en esta propuesta: *procesos psicológicos subyacentes, procesos implicados en la asociación de grafemas y fonemas, procesos implicados en la integración textual, aquellos relacionados con la integración texto-conocimientos y experiencias previas, procesos metacognitivos (procesos de autorregulación y metaconocimientos) y procesos de ajuste personal-social*; así como los indicadores que permitirán a los evaluadores valorar su presencia o puesta en práctica. Igualmente, se recoge en este capítulo un esquema de la totalidad de las actividades de evaluación-intervención propuestas, así como la red de relaciones que guardan estas actividades en función de los procesos que evalúan o sobre los que actúan, y también del camino diseñado para favorecer un uso más eficaz —a la vez que flexible— del conjunto de actividades.

En el capítulo siguiente se desarrollan todas las actividades de evaluación-intervención que forman parte de la propuesta, en las que se explicitan (a) los procesos que evalúa cada actividad, (b) la descripción detallada de la misma, (c) la metodología propuesta para su aplicación, (d) los patrones metodológicos de actuación relacionados con la

mediación y el análisis de procesos metacognitivos, lo que nos permitirá obtener información útil sobre los procesos de aprendizaje del alumno y las dificultades encontradas en el acceso y desarrollo de la lectura, para poder realizar, posteriormente, propuestas de intervención más ajustadas, y por último, se detallan (e) los criterios de evaluación, que permitirán y facilitarán valorar el proceso de resolución de la tarea y los resultados obtenidos. En el capítulo nueve se detalla nuestra propuesta específica de intervención para la mejora de los procesos metacognitivos en la lectura. Esta propuesta se basa en la repetición y análisis exhaustivo —siguiendo una guía de cuestiones a realizar— de las actividades de evaluación-intervención pertenecientes a aquellos procesos en los que el alumno ha manifestado problemas. También en este capítulo se recogen algunas referencias a otras propuestas de intervención que guardan relación con la nuestra. Por último, ya en el capítulo de anexos, se encuentran, por un lado, las tareas concretas que el alumno debe realizar y que nos servirán para desarrollar, durante su puesta en práctica, nuestro sistema de evaluación dinámica; y por otro, en la segunda parte de ese mismo capítulo encontraremos las Hojas de registro y evaluación de cada una de las actividades. Estas hojas de registro y evaluación permitirán recoger de un modo sencillo, práctico y atendiendo sobre todo a aspectos cualitativos, la información que nos parezca relevante para poder llegar a mejorar los procesos evaluados.

Este trabajo pretende ser una herramienta útil y eficaz para la labor de evaluación, intervención y asesoramiento psicopedagógico relacionada con las dificultades de aprendizaje en la lectura y que habitualmente desarrollan orientadores, profesorado de apoyo a la integración y profesorado de audición y lenguaje en los Centros escolares de Educación Primaria y Secundaria. Sin duda, esta propuesta cobra sentido si esta labor se realiza en modo colaborativo con el profesorado del alumno a quien pretendemos ayudar. También puede ser utilizado por los mismos profesores del alumno o alumnos que presenta(n) dificultades. La utilización de este dispositivo de evaluación-intervención puede ser enormemente flexible, de modo que, en función de las necesidades educativas especiales que se presenten en un aula o asociadas a un alumno en concreto, pueden llevarse a cabo determinadas actividades pertenecientes a un solo bloque de procesos o poner en práctica algunos patrones metodológicos de actuación de los que aquí se proponen.

Por otra parte, durante la fase de construcción y reconstrucción del dispositivo de evaluación que ahora presentamos, se ha recogido de forma continuada información de su puesta en práctica en modo experimental. Esta ha sido llevada a cabo por sus autores con el objeto de valorar la utilidad y eficacia de las actividades que forman parte del dispositivo, así como de los patrones metodológicos propuestos. Una vez finalizado, se ha procedido a poner en práctica el dispositivo a modo de aplicación piloto con 9 alumnos/as, la mayoría de ellos en Educación Primaria (desde 2º hasta 6º curso). Esta aplicación piloto tuvo lugar durante el curso escolar 2001/2002 en dos centros públicos, uno situado en Arahal y otro en Alcalá de Guadaíra (Sevilla). En todos los casos se ha tratado de alumnos con especiales dificultades de aprendizaje de la lectura que eran atendidos a tiempo parcial en aulas de apoyo a la integración o que, de alguna forma, recibían refuerzo educativo por parte de otros profesores del centro. Las dificultades versaban principalmente sobre el dominio del código en condiciones que facilitarían el acceso a la comprensión de los textos, es decir, problemas fundamentalmente en la decodificación (4 alumnos); sobre los problemas en la comprensión global de los textos, sin que existieran notables dificultades en el acceso al código (3 alumnos); y también formaban parte de esta reducida muestra dos sujetos con dificultades tanto en la

decodificación como en la comprensión global de los textos trabajados. Básicamente la información sobre el proceso y sobre los resultados obtenidos concuerdan con los ya expuestos anteriormente a raíz de la investigación sobre evaluación dinámica de metacogniciones (Navarro y Mora, 2003b), con la salvedad de que, en este caso, se trataba de valorar, además de los procesos metacognitivos (metacogniciones y procesos de autorregulación) implicados en la comprensión, otros procesos también relevantes en la explicación de las dificultades de aprendizaje.

Hemos constatado que niños y niñas que presentan serias dificultades en los procesos de decodificación, que se manifestaban en una lectura apegada al texto, excesivamente centrada en la decodificación «correcta» de las palabras y que se relacionaba, a su vez, con bajos índices de comprensión global de los textos trabajados, mejoraban su rendimiento de forma significativa cuando la mediación del profesor/a se dirigía a llevar su atención sobre habilidades de control y regulación consciente de su propia comprensión. Indicaciones dadas por el profesor encaminadas a hacer pensar al alumno sobre lo que estaba leyendo, a detenerse si no comprendía algo, y a pensar y valorar al final de su lectura si había comprendido, parecieron contribuir notablemente a mejorar las realizaciones de los alumnos evaluados en las actividades propuestas. En general, el apoyo y la mediación ofrecidos han estado encaminados a ofrecer estas herramientas cognitivas de control y regulación del propio proceso de lectura y comprensión; ha sido, por tanto, fundamentalmente un apoyo de corte metacognitivo.

Hemos constatado también que todos los alumnos evaluados mostraban deficiencias precisamente en los procesos metacognitivos. Algunos, sobre todo los más pequeños, mostraron un bajo nivel de metacognición estratégico y sobre el significado y sentido de la lectura; pero, en general, todos mostraron de partida un bajo nivel de autorregulación consciente del proceso de comprensión y, por tanto, se mostraban poco hábiles para poner en práctica habilidades de planificación de su lectura, de supervisión continua de su propia comprensión y de evaluación de la misma.

Se pusieron en práctica, a lo largo de esta fase de construcción y aplicación piloto del dispositivo, todas las actividades elaboradas y recogidas en nuestra propuesta de evaluación dinámica de procesos implicados en la lectura. Nuestra experiencia, sin embargo, al afrontar la evaluación de los nueve alumnos citados anteriormente, ha sido que, en general, no es necesario realizarlas todas. En la mayoría de los casos, dependiendo de la demanda inicial de otros profesores o del problema inicial detectado por nosotros mismos, han bastado 5 ó 6 tareas, a raíz de las cuales se ha extraído información que se ha considerado idónea y relevante para plantear una intervención orientada a la mejora de los problemas. Así mismo, en cuanto al tiempo empleado, éste ha oscilado entre las cuatro y las seis sesiones como media, incluyendo la entrevista con el profesor tutor del alumno en dificultades.

Habitualmente, aún cuando muchos de los alumnos que hemos evaluado presentaban inicialmente dificultades manifiestas relacionadas con los procesos de asociación grafema-fonema, hemos comenzado con la lectura de un pequeño texto, casi siempre perteneciente a una de las actividades propuestas. La lectura de este pequeño texto ofrecía ya abundante información al evaluador preparado, que estaba sumamente atento y sensibilizado para encontrar en el desarrollo de la actividad, elementos que revelasen, por ejemplo: la existencia de dificultades, estrategias que ponía el alumno en marcha de forma espontánea, actitudes mostradas hacia las tareas de lectura, conocimientos sobre

los contenidos, relaciones que establecía, dudas que manifestaba, etc. Hemos aprovechado, también desde un primer momento, para valorar especialmente sus metacnocimientos sobre estrategias de comprensión, así como sobre el significado y sentido de la lectura. Posteriormente, si se ha considerado oportuno, se ha trabajado con otras actividades propuestas en el dispositivo que posteriormente presentamos; habitualmente: *lectura de pseudo-palabras*, *conciencia fonológica*, *integración texto-conocimientos previos y realización de inferencias*, así como con las tres actividades pertenecientes al bloque de evaluación de procesos metacognitivos. En este mismo capítulo tendremos oportunidad de visualizar y profundizar en el esquema que, en modo orientativo, ofrecemos para la puesta en práctica del dispositivo.

Los procesos implicados en la comprensión de textos que han sido contemplados en el dispositivo

Sobre los distintos procesos que intervienen en la lectura existe ya una abundante información en la literatura científica especializada (Repetto, Téllez y Beltrán, 2002; Sánchez, 1999; Cuetos, Rodríguez y Ruano, 1996; Alonso Tapia y Cols., 1992). Nosotros mismos hemos hecho referencia, a lo largo del primer capítulo de este trabajo, a numerosos estudios y autores que inciden en la relevancia de intervenir y evaluar determinados procesos (lingüísticos, psicolingüísticos, cognitivos, socio-personales, metacognitivos...) asociados de una u otra forma a la lectura. Por otra parte, se ha realizado un análisis, tanto de los procesos mayormente citados en la literatura y de algunos otros que hasta ahora no han sido suficientemente tenidos en cuenta, como de las tareas que mejor responderían a una evaluación del proceso analizado. Este análisis ha dado como resultado la contemplación en nuestra propuesta de siete tipos de procesos implicados en la lectura, que hemos agrupado a su vez en tres grandes bloques: *Procesos Metacognitivos*, que incluye: (a) los procesos de conocimiento consciente sobre la lectura y sus contenidos, sobre las estrategias que pueden aplicarse (metacnocimiento estratégico), y sobre los procesos de ajuste personal-social; y (b) los procesos de autorregulación de la comprensión; *Análisis e Integración de la Información*, que incluye: (c) los procesos implicados en la asociación de grafemas y fonemas; (d) los procesos psicológicos subyacentes; (e) los procesos implicados en la integración textual; y (f) los procesos de integración texto-conocimientos previos; y, por último, los *Procesos implicados en el Ajuste Personal-Social*. En total se contemplan 42 procesos, que aparecen recogidos en el cuadro 29.

Ya hemos hecho mención aquí acerca de la relevancia, en la evaluación de la lectura, de los procesos metacognitivos, tanto de los metacnocimientos como de los procesos de autorregulación. También existe un amplio consenso sobre la conveniencia de evaluar procesos implicados en la integración textual y procesos implicados en la integración de la información obtenida con los conocimientos y experiencias previas, para que así se produzcan aprendizajes verdaderamente significativos para el alumno, se pueda revisar el contenido del texto desde lo que ya se conocía y se pueda acceder a la valoración de la información (Solè, 1992; Carriedo y Alonso, 1994; García Madruga, Martín, Luque y Santamaría, 1995; Sánchez, 1998; 1999). Igualmente hemos hecho referencia anteriormente a la importancia que, según numerosas investigaciones, tienen los procesos implicados en la asociación de grafemas y fonemas para la eficacia lectora de los alumnos y alumnas. Nos queda justificar en alguna medida la presencia en nuestro dispositivo de evaluación, de los procesos psicológicos subyacentes y de aquellos procesos implicados en el ajuste personal-social. En cuanto a los primeros, decir que,

aunque es cierto que según las investigaciones, pocos niños presentan dificultades de lectura asociadas a problemas en percepción visual, coordinación visual, discriminación perceptiva o lateralidad, y que, por otra parte, no se han encontrado evidencias de causalidad entre trastornos viso-espaciales y dificultades de lectura (Alegría, 1985; Ellis, 1984); estos pocos niños, así como la presencia de este tipo de procesos, ya sea en forma de prerrequisitos necesarios para la lectura o en forma de procesos subyacentes que se encuentran presentes siempre que se accede a la lectura, justifican el hecho de que tratemos de evaluar su incidencia en la lectura. De este modo también nos servirá para descartar, en su caso, la presencia de este tipo de dificultades cuando exista alguna sospecha de que pueden encontrarse ahí los problemas.

En cuanto a los problemas relacionados con factores emocionales y de ajuste personal-social, es necesario destacar la relevancia y la interconexión que presentan varios de estos aspectos. En este sentido, la presencia en determinados alumnos de sentimientos de frustración, incapacidad o fracaso, unidos a menudo a una percepción devaluada de sí mismo y de sus aptitudes y posibilidades académicas, así como al posible efecto de reforzamiento recíproco entre las pobres expectativas, por una parte, y (a) la percepción negativa del profesorado acerca de las posibilidades de aprendizaje del alumnado con dificultades de aprendizaje, o (b) la autopercepción negativa que los propios alumnos tienen sobre su capacidad y sobre las posibilidades de éxito en sus actividades, por otra, (Meltzer, Reddy, Pollica, Roditi, Sayer y Theokas, 2004; Meltzer, Katzir, Miller, Reddy y Roditi, 2004); suele desembocar en actitudes de rechazo hacia la lectura y hacia las actividades encaminadas a su aprendizaje. La consideración de la importancia de estos aspectos para el desarrollo educativo y socio-personal de los sujetos, así como su compleja interdependencia, justifican sobradamente la inclusión de este tipo de factores en cualquier proceso de evaluación de la lectura, siendo necesario, no sólo tenerlos en cuenta, sino analizar su influencia y evaluarlos (Romero, 1990; Mora, 1991; Moreno y García, 2000; Sánchez, 1999; Mora-Merchán y Mora, 2000). Por otra parte, la evaluación del contexto de aula, donde se ponen en práctica habitualmente los procesos de enseñanza y aprendizaje de la lectura, cobra especial importancia —que ya ha sido más ampliamente tratada con anterioridad— tanto en la evaluación de las dificultades asociadas a la lectura, como en la determinación de las condiciones favorecedoras de un mejor desarrollo de estos procesos (agrupamientos, tipo de apoyo más idóneo para el alumno, modo de acometer la enseñanza de la lectura o el desarrollo de habilidades específicas, etc.).

Indicadores de evaluación de los procesos implicados

La evaluación de la lectura que proponemos pretende dar cuenta del análisis de los distintos procesos implicados en la actividad lectora. Para esto, como ya hemos comentado, se hace necesario bajo nuestro punto de vista, evaluar *el proceso de lectura*, es decir, evaluar en el momento en el cual se está leyendo y evaluar precisamente el proceso de resolución de la tarea de lectura, y no tanto el resultado de la misma. Esta evaluación de procesos conlleva ciertas dificultades, derivadas mayormente de la «invisibilidad» de los procesos cognitivos y emocionales puestos en juego, debiendo realizar el evaluador continuas inferencias acerca de las habilidades que se emplean en las tareas propuestas. Para tratar de paliar, en la medida de lo posible estas dificultades, se ha elaborado un listado de indicadores —o criterios de evaluación— que faciliten la tarea de determinar la presencia o no de los procesos implicados en la lectura. Se trata, en cierta forma, de *descomponer* el complejo proceso de lectura en las habilidades

Cuadro 29. Procesos y elementos implicados en la lectura

Procesos Metacognitivos		Análisis e Integración de la Información				Procesos implicados en el ajuste personal-social
Metaconocimientos sobre lectura y sus contenidos, estrategias de lectura y ajuste P-S	Procesos implicados en la autorregulación de la comprensión	Procesos implicados en la asociación grafema-fonema	Procesos psicológicos subyacentes	Procesos implicados en la integración textual	Procesos implicados en la integración texto-conocimientos	
<p><i>Significado y sentido de la lectura (en qué consiste leer y para qué sirve)</i></p> <p><i>Estrategias propias que pueden aplicarse</i></p> <p><i>Estructura fonológica, léxica y sintáctica del lenguaje</i></p> <p><i>Estructuras textuales y géneros lingüísticos</i></p> <p><i>Adecuación del mensaje en función del contexto</i></p> <p><i>Contenidos de los textos</i></p> <p><i>Actitudes, motivación e intereses</i></p> <p><i>Dificultades de aprendizaje, expectativas, autoconcepto y vías de solución</i></p>	<p><i>Actividades de planificación</i></p> <p><i>Actividades de supervisión durante la lectura</i></p> <p><i>Actividades de autoevaluación</i></p> <p><i>Transferencia y Generalización</i></p>	<p><i>Identificación/reconocimiento de fonemas y grafemas</i></p> <p><i>Unión-segmentación fonémica y silábica</i></p> <p><i>Asociación grafemas-fonemas</i></p>	<p><i>Percepción y discriminación visual</i></p> <p><i>Atención</i></p> <p><i>Coordinación visual</i></p> <p><i>Estructuras rítmicas</i></p> <p><i>Orientación espacial y temporal</i></p> <p><i>Memoria</i></p>	<p><i>Integración de todas las ideas en un esquema</i></p> <p><i>Construcción de proposiciones globales</i></p> <p><i>Integración lineal de las ideas</i></p> <p><i>Construcción de proposiciones</i></p> <p><i>Distinción de la información relevante</i></p> <p><i>Jerarquización y secuenciación de la información</i></p> <p><i>Acceso a la información contenida en cuadros y esquemas</i></p> <p><i>Identificación y uso de la estructura textual</i></p> <p><i>Acceso al significado de palabras</i></p>	<p><i>Activación de los conocimientos necesarios</i></p> <p><i>Integración de nuevas ideas con conocimientos y experiencias anteriores</i></p> <p><i>Establecimiento de relaciones</i></p> <p><i>Formulación y respuesta de preguntas acerca del texto con relación a sus conocimientos previos</i></p> <p><i>Formulación de hipótesis y anticipación de resultados</i></p> <p><i>Realización de inferencias</i></p> <p><i>Revisión de las ideas del texto desde lo que ya se sabe</i></p>	<p><i>Motivación, actitudes e interés con relación a la lectura</i></p> <p><i>Cooperación con los/as compañeros/as en tareas de lectura</i></p> <p><i>Conexión de las actividades de lectura con el currículum</i></p> <p><i>Experiencias con la lectura fuera de la escuela</i></p> <p><i>Procesos de enseñanza y aprendizaje de la lectura en el aula</i></p>

necesarias para llevarlo a cabo, con el objeto de facilitar su análisis. Estos indicadores, obviamente, tratan de evaluar los procesos detallados anteriormente en el cuadro 29 y se han redactado de la forma más clara y concreta posible para facilitar su identificación y valoración. A continuación exponemos los indicadores de evaluación para cada uno de los procesos implicados en la lectura.

Cuadro 30. Procesos implicados en la lectura e indicadores para su análisis

Procesos Metacognitivos
<p>INDICADORES DE METACONOCIMIENTO SOBRE LA LECTURA Y SUS CONTENIDOS, ESTRATEGIAS DE LECTURA Y AJUSTE PERSONAL-SOCIAL</p> <p>Es consciente de que leer sirve para obtener información Es consciente de que leer significa comprender lo que se ha leído Es consciente de la estructura fonológica, léxica y sintáctica del lenguaje Es consciente de lo que sabe sobre elementos estructurales y tipologías de textos Conoce las propias estrategias que puede aplicar antes, durante y después de leer Conoce cómo y cuándo se aplican las propias estrategias y por qué son útiles Es consciente de lo que sabe acerca del contenido de la lectura Es consciente de lo correcto/incorrecto de un mensaje, desde un punto de vista semántico, en función del contexto Es consciente de su propia actitud, motivación e intereses relacionados con la lectura Es consciente de las propias DA relacionadas con la lectura, de sus expectativas y autoconcepto, así como de las posibles vías de solución a adoptar</p>
<p>INDICADORES DE PROCESOS IMPLICADOS EN LA AUTORREGULACIÓN DE LA COMPRENSIÓN</p> <p><i>Actividades de planificación</i></p> <p>Dirige por sí mismo/a procesos de reflexión y solución del problema planteado Identifica la tarea que debe realizar y el problema planteado Selecciona los recursos más adecuados para realizar la tarea Define y crea metas Anticipa errores Selecciona los recursos más adecuados para cada error Define/establece prioridades Secuencia operaciones a realizar</p> <p><i>Actividades de supervisión durante la lectura</i></p> <p>Realiza actividades coherentes con la meta que ha definido, manteniendo los objetivos Manifiesta constancia de estrategias y de la secuencia de operaciones Supera obstáculos de la tarea Manifiesta flexibilidad en el uso de planes y estrategias Manifiesta control de la impulsividad. Presenta latencias Gestiona los errores positivamente Destaca los absurdos y las incoherencias lingüísticas Utiliza estrategias de memoria Supervisa sin necesidad de mediación externa constante Utiliza la mediación verbal durante la tarea. Se cuenta a sí mismo lo que va a hacer, lo que está haciendo o lo que ha hecho, se formula preguntas, se plantea dudas, etc. Se recupera de las interrupciones y distracciones Distribuye la atención de forma que se centra en los contenidos relevantes Controla la actividad durante la lectura para determinar si entiende o no lo que está leyendo</p>

Actividades de autoevaluación

Evalúa logros con relación a las metas definidas
 Evalúa la adecuación de la conducta con relación a la consecución de las metas
 Evalúa la adecuación de las metas
 Evalúa la adecuación de las estrategias
 Formula conclusiones autoalusivas
 Verifica la información obtenida de la lectura
 Se formula preguntas y dudas sobre el texto leído

Transferencia y Generalización

Formula/Reconoce principios o leyes de carácter general
 Aplica estrategias o conclusiones a/de otras tareas

Procesos de Análisis e Integración de la información**INDICADORES DE PROCESOS IMPLICADOS EN LA ASOCIACIÓN G-F**

Identifica/reconoce los fonemas y grafemas
 Realiza operaciones de unión de fonemas y sílabas para formar una palabra
 Realiza operaciones de segmentación de una palabra. Es capaz de discriminar/aislar los fonemas
 Realiza operaciones de asociación grafema-fonema y fonema-grafema

INDICADORES DE PROCESOS PSICOLÓGICOS SUBYACENTES

Realiza adecuadamente el seguimiento visual de un texto. No se salta palabras ni líneas
 Identifica dibujos o símbolos lingüísticos/no lingüísticos con igual/diferente orientación espacial
 Identifica/reconoce y discrimina los grafemas y/o conjunto de grafemas visualmente
 Mantiene la atención durante la realización de la tarea
 Estructura temporalmente de forma adecuada una serie de sucesos
 Almacena, recupera y utiliza información de manera simultánea a la realización de la tarea
 Almacena y recupera información a corto plazo, que utiliza en las tareas propuestas
 Almacena y recupera información a largo plazo, que utiliza en las tareas propuestas

INDICADORES DE PROCESOS IMPLICADOS EN LA INTEGRACIÓN TEXTUAL

Identifica/reconoce y accede al significado de palabras escritas
 Estructura de forma adecuada una oración
 Estructura adecuadamente el mensaje atendiendo a los signos de puntuación
 Construye el significado de proposiciones
 Vincula proposiciones y construye significados aplicando reglas de supresión y generalización
 Integra adecuadamente palabras conectivas o funcionales (preposiciones, conjunciones, adverbios y determinantes)
 Identifica y accede a la información contenida en cuadros y esquemas
 Estructura el resumen del texto de forma adecuada, realizando un texto con sentido y estructura definidas
 Resume el texto con un sentido global y utilizando vocabulario propio
 Sigue la progresión temática adecuadamente al realizar o identificar un resumen del texto
 Jerarquiza y secuencia la información
 Justifica y razona respuestas y decisiones
 Identifica la estructura del texto. Utiliza elementos estructurales para comprender el texto
 Identifica ideas relevantes de un texto

INDICADORES DE PROCESOS IMPLICADOS EN LA INTEGRACIÓN TEXTO-CONOCIMIENTOS

Integra nuevas ideas con conocimientos y experiencias anteriores
Utiliza conocimientos y experiencias anteriores para construir significados
Es sensible al estilo, los presupuestos, la perspectiva y las pretensiones del autor/a y hasta puede llegar a cuestionar estos aspectos
Expresa opiniones, juicios o intuiciones acerca de los contenidos del texto
Establece relaciones entre diversos textos
Formula y responde preguntas acerca del texto
Formula hipótesis. Anticipa resultados
Contrasta y verifica las hipótesis
Realiza inferencias relativas al texto y que lo sobrepasan
Revisa las ideas del texto desde lo que ya sabe

Procesos de Ajuste personal-social

INDICADORES DE PROCESOS IMPLICADOS EN EL AJUSTE PERSONAL-SOCIAL CON RELACIÓN A LA LECTURA

Motivación, actitudes e interés con relación a la lectura

Los objetivos son intrínsecos y orientados hacia el dominio de la lectura
Persiste cuando se enfrenta a obstáculos o dificultades
Elige tareas retadoras al límite de su capacidad actual
Muestra entusiasmo por la lectura/Demuestra interés durante el desarrollo de actividades relacionadas con la lectura
Exhibe orgullo y confianza en sí mismo/a como lector/a
Considera que puede mejorar su propia lecto-escritura. Se asigna un papel activo
Considera que los demás respetan sus aportaciones
Corrige los propios errores sin agresión o depresión. No manifiesta ansiedad o miedo al fracaso
Mantiene opiniones propias, cuando ello esté justificado. No cede a la presión social y contrarresta las contrasugerencias
Selecciona voluntariamente la lectura y la escritura como actividades de opción libre
Elige adecuadamente textos para leer y temas sobre los que escribir (a juicio del evaluador/es)

Cooperación con los/as compañeros/as en tareas de lectura

Frecuentemente colabora en actividades de lectura
Inicia o participa activamente en discusiones, diálogos o debates sobre significados de textos
Ofrece apoyo positivo, afecto y soporte educativo a sus compañeros/as
Desempeña una variedad de roles en la comunidad de estudio
Valora las aportaciones de los demás, respeta las opiniones y la ayuda ajena
Aparición/petición de conductas de ayuda/colaboración

Conexión de la lectura con el currículum

Comprende que la lectura es un instrumento para aprender y para conocerse a sí mismo/a
Contempla la lectura, la escritura, el habla y la escucha como actividades que se apoyan mutuamente
Entiende que lo que se aprende al leer y escribir es útil en otras materias

Experiencias en lecto-escritura fuera de la escuela

La lectura y la escritura forman parte de sus actividades rutinarias cotidianas
Conecta las actividades escolares de lecto-escritura con la lectura y escritura en la vida cotidiana
Se siente animado y apoyado para leer y escribir fuera de la escuela
En su hogar existe un ambiente estimulador hacia la lectura y la escritura

Procesos de enseñanza y aprendizaje en el aula

Se fomenta la activación de conocimientos y experiencias previas, y su relación con la lectura
 Se potencian las actividades de planificación de la lectura
 Se fomenta la supervisión durante la lectura
 Se favorecen actividades de autoevaluación en los alumnos/as
 Se facilita la discusión y la reflexión sobre los contenidos de la lectura, los distintos elementos que la conforman, sobre las propias estrategias y sobre el proceso de lectura
 Se enseña y/o facilita el uso de estrategias de extracción, integración y construcción de significados
 Se evalúan procesos implicados en la lectura, se analizan las dificultades encontradas y qué tipo de ayuda sería útil para superar esas dificultades e interiorizar estrategias de lectura

El registro de la evaluación del proceso de resolución de las actividades y de los resultados obtenidos

En el capítulo de anexos del material que conforma el dispositivo EDPL y que fue proporcionado a los profesores y orientadores que han colaborado en la investigación, se podía encontrar, además de las tareas concretas que debía realizar el alumnado participante, las *Hojas de registro y evaluación* de cada una de las actividades de evaluación-intervención propuestas. En éstas se valoran los resultados obtenidos por el alumno al realizar la tarea, pero sobre todo sirven al propósito de valorar y recoger el *proceso de resolución de la actividad, las posibles dificultades encontradas, la mediación efectuada y en qué medida se ha beneficiado el alumno de esta mediación incorporando algunas de las pautas o guías que le ofrecíamos, y, por último, también se evalúa la puesta en práctica de estrategias metacognitivas*. Para facilitar la tarea de evaluación del proceso de resolución de cada una de las actividades, los aplicadores disponían de los indicadores de la puesta en práctica de los procesos implicados para cada una de las tareas; estos indicadores permitían tener presente los criterios de evaluación de cada actividad, pudiendo marcar aquellos que el alumno había puesto en práctica en la tarea que se estuviera realizando y realizar las observaciones consideradas oportunas.

En el dispositivo se realiza un énfasis especial en la relevancia de recoger y valorar el proceso de mediación efectuado con el fin de obtener información útil sobre los procesos de aprendizaje del alumnado, las dificultades encontradas, así como especialmente sobre las *actuaciones* desarrolladas por el evaluador que han permitido facilitar de forma significativa la realización de la tarea por parte del alumno, accediendo de esta forma a *zonas de desarrollo potencial* en las que nuestra actividad de evaluación haya cobrado plenamente sentido. Igualmente relevante, debido a las mismas razones expuestas anteriormente, se considera la valoración más cualitativa del proceso de resolución de la actividad, así como de la puesta en práctica de los procesos metacognitivos. Por último, el análisis posterior y conjunto entre evaluador y profesor del alumno, en el marco de un proceso de evaluación al margen de esta aplicación experimental del dispositivo, debería facilitar el acceso a un mayor conocimiento de las dificultades concretas que presenta el alumno y de cómo incidir en la mejora del proceso de lectura.

Por otra parte, en otro de los anexos que acompaña al material original se encuentra una Hoja de registro y evaluación en la que se recoge la totalidad de los indicadores. Esta ha

sido confeccionada para valorar la presencia o ausencia de los procesos implicados en la lectura, así como para realizar aquellas observaciones que nos parezcan relevantes a la hora de determinar necesidades. Es una plantilla que permite sobre todo —al igual que las Hojas de registro y evaluación de cada una de las tareas— un análisis cualitativo de la información valorada, donde se recogen todos los indicadores propuestos para la evaluación, con lo que, una vez rellena, obtendremos una visión y valoración globales acerca de: (a) las dificultades existentes, (b) dónde se encuentran principalmente estas dificultades, (c) qué aspectos contextuales están influyendo, (d) en qué medida se beneficia el alumno evaluado de nuestra mediación y apoyo en las tareas, etc. Esta Hoja de registro y evaluación que contiene todos los indicadores, puede ser de gran utilidad al final del proceso de evaluación, ya que al permitir una visión global de la evaluación de cada una de las habilidades puestas en práctica por parte del alumno y de las posibles dificultades encontradas, facilitará la síntesis evaluativa, así como las propuestas de intervención.

Esquema general de la Evaluación dinámica de procesos implicados en la lectura y conexiones entre las actividades de evaluación propuestas

Esquema de las actividades de evaluación

A continuación exponemos en el cuadro 31 el esquema de las actividades de evaluación que forman parte del dispositivo y a través de las cuales se valoran los indicadores recogidos anteriormente. Como queda anotado en el apartado *Contextualizar los instrumentos de evaluación*, del capítulo 3, estas actividades pueden ser completadas o incluso sustituidas por otras que sean construidas por los propios evaluadores y que se ajusten mejor a las características del alumnado evaluado.

Cuadro 31. Esquema de las actividades de evaluación

Procesos Psicológicos Subyacentes

- Tareas de Coordinación y Seguimiento visual
 - Lectura de un texto
 - Laberinto
- Discriminación perceptiva/ Orientación espacial
 - Discriminación de letras en palabras
 - Discriminación de letras entre otras dadas
 - Comparación y discriminación de símbolos no lingüísticos
- Emparejamiento de palabras iguales
- Composición de palabras con letras sueltas
- Secuenciación temporal
- Memoria

Procesos implicados en la Asociación Grafema-fonema

- Lectura de Pseudopalabras

-
- Evaluación de la conciencia fonológica
 - Decodificación fonológica y asociación grafema-fonema con pseudopalabras

Procesos de Integración textual

- Lectura de palabras

-
- Lectura de palabras apoyadas con dibujos
 - Lectura de pseudohomófonos y homófonos
 - Incorporación de palabras al léxico interno visual

- | | | | | |
|---|--------------|---|--------------|--|
| <ul style="list-style-type: none"> <input type="checkbox"/> Ordenar frases <input type="checkbox"/> Relacionar frases—
dibujos <input type="checkbox"/> Elaborar frases nuevas
a raíz de otras dadas | <p>.....</p> | <ul style="list-style-type: none"> - Completar oraciones con
la palabra que falte... | <p>.....</p> | <ul style="list-style-type: none"> - Construcción oral de
oraciones |
|---|--------------|---|--------------|--|

- Lectura en voz alta de un texto para evaluar signos de puntuación/ritmo
- Uso de la estructura del texto
- Extracción y Construcción del significado

Integración texto-conocimiento

- Integración texto-conocimientos previos y realización de inferencias
- Elaboración del final de un texto
- Conocimientos y experiencias previas
- Establecimiento de relaciones
- Elaboración de hipótesis

Procesos metacognitivos

- Incoherencias lingüísticas
- Comparar dos textos
- ¿Qué podemos hacer para comprender mejor?

Ajuste personal-social

- Evaluación del contexto

Esquema del desarrollo de las actividades de evaluación

En el desarrollo de cada una de las actividades de evaluación propuestas, se recogen los siguientes aspectos:

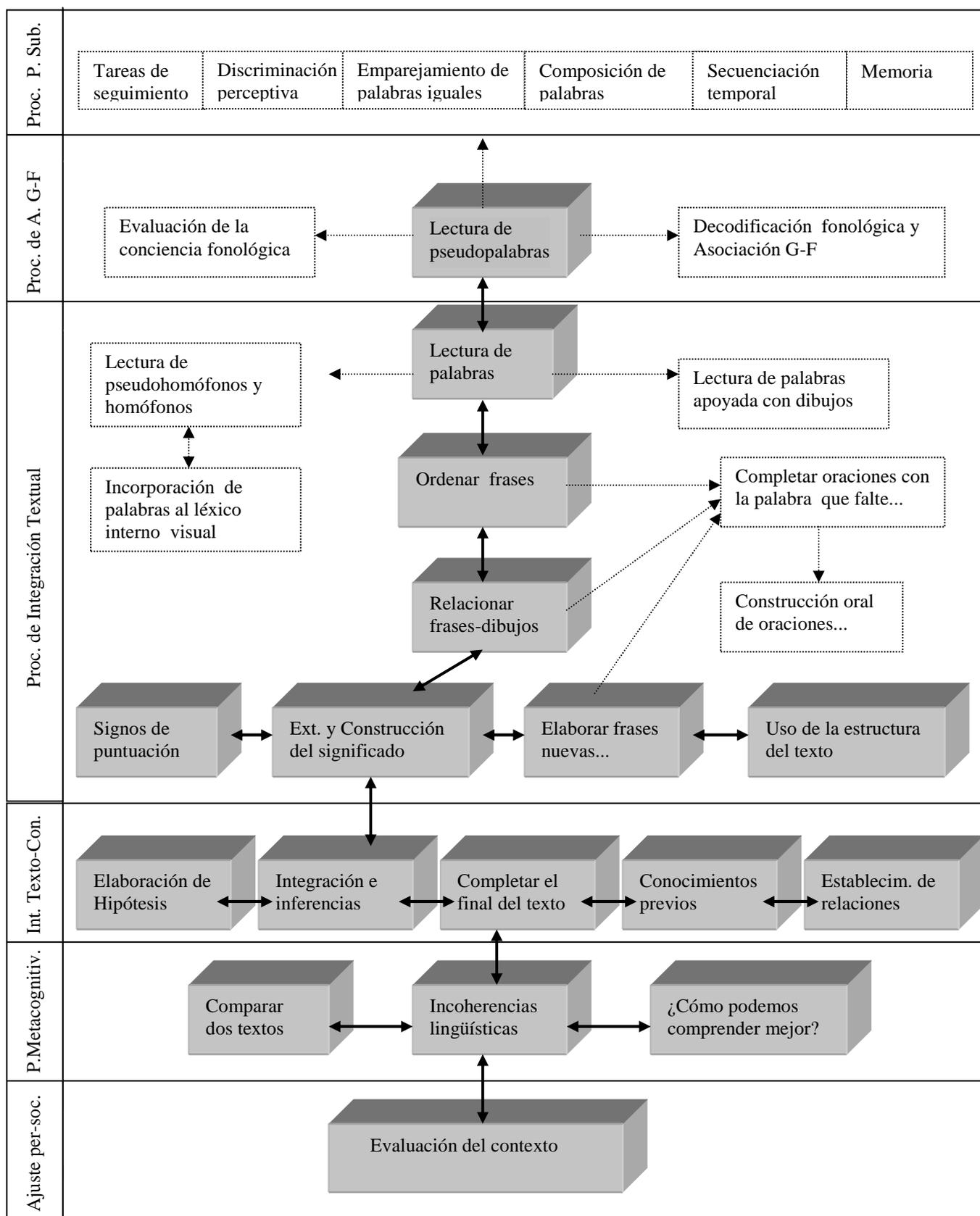
- *Proceso que se pretende evaluar.* Tomando como referencia el cuadro 29, se detallan al comienzo de cada actividad el o los procesos principales que se quieren valorar mediante la realización de la actividad. También se recogen otros procesos que pueden valorarse en la tarea. Habitualmente, entre estos procesos se encontrarán las habilidades metacognitivas, ya que nuestra intención será valorar su puesta en práctica por parte del alumno durante la realización de cada una de las actividades.
- *Descripción de la actividad de evaluación.* En este apartado se describe el contenido de la actividad, especificando los elementos de los que consta, referencias que creemos pueden ser de interés para el lector, así como el contenido de los textos, frases o palabras, sobre los que el alumno evaluado deberá trabajar.
- *Metodología propuesta para su aplicación.* Se recoge una propuesta metodológica para llevar a cabo la actividad. En este apartado se detalla de forma sucinta cómo se presenta la tarea al alumno y qué explicaciones iniciales serían necesarias.
- *Mediación y evaluación de procesos metacognitivos.* Aquí se detalla nuestra propuesta metodológica para la evaluación preferente de los procesos de control y regulación consciente de la lectura. Se expone, en cada una de las actividades, el *proceso de mediación* que puede utilizar el profesor o el evaluador para acceder de una forma más segura y clara al proceso de lectura del alumno evaluado, observar y analizar las posibles dificultades que éste encuentra en el *camino* de resolución de la actividad y poder establecer así cuál es el tipo y grado de apoyo que necesita. Este proceso de mediación recoge principalmente distintas cuestiones que el evaluador puede hacer al alumno mientras realiza la actividad o una vez finalizada ésta, con la intención de valorar el proceso mismo de resolución de la tarea y también en qué medida se beneficia de la mediación realizada y del apoyo ofrecido durante la actividad. También se detallan los procesos de lectura que podemos *observar* de forma prioritaria en la actividad, a los que debemos prestar una especial atención. Es en este apartado, por tanto, donde se ponen en práctica dos de los elementos que nos servían al propósito de evaluar las habilidades de autorregulación y los metaconocimientos en cada una de las actividades propuestas: por una parte, la entrevista metacognitiva y, por otra, la observación sistemática.
- *Criterios de Evaluación.* Tomando como referencia los indicadores de evaluación que se encuentran en el cuadro 30, se recogen al final de cada actividad los criterios de evaluación que se corresponden con los procesos principales que nos proponíamos evaluar con esa actividad, así como otros indicadores que igualmente podemos evaluar y que, a su vez, se corresponden con los otros procesos que también podíamos valorar con la actividad y que se recogen al comienzo de la misma.

Galaxia de las actividades de evaluación

Aunque las actividades de evaluación que se proponen, admiten, como ya hemos comentado, un alto grado de flexibilidad en cuanto al contenido que se dé a la actividad en sí, en función de las características del alumnado evaluado y del contexto donde la evaluación tiene lugar, así como en cuanto a la secuencia que podemos seguir a la hora de evaluar los procesos de lectura; proponemos en la página siguiente (figura 4) una *galaxia* de las actividades de evaluación, citadas anteriormente, en la que se puede observar la red de relaciones que mantienen unos instrumentos con otros en función de la secuencia que consideramos idónea a priori. Las actividades que se encuentran en rectángulos tridimensionales y sombreados permiten evaluar la lectura atendiendo a aquellos procesos que son considerados más relevantes para explicar las posibles dificultades. Es posible, por tanto, que ante evaluaciones efectuadas con un carácter de investigación, descriptiva o explicativa, se opte por valorar el proceso de resolución de todas estas actividades, o de algunas de ellas, al margen de valorar también las otras si se cree oportuno. De otro modo, en evaluaciones con carácter diagnóstico y de posterior intervención en dificultades en la lectura, sugerimos que se ponga en práctica la flexibilidad de la que hemos hecho mención anteriormente, es decir, podemos comenzar a evaluar utilizando aquellas actividades que mejor se adecuan a las dificultades que ha manifestado el alumno hasta entonces, e ir acotando dónde se encuentran sus principales problemas, a la vez que vamos determinando algunas condiciones que mejorarían su rendimiento, «deslizándonos» por el resto de actividades que se relacionan con el problema presentado y que nos ayudarán a determinar más concretamente el problema, así como las posibles soluciones. Un ejemplo puede clarificarnos este punto: ante un problema detectado en un niño que lee con dificultad, silabea continuamente y no da muestras de comprender la mayor parte de lo que ha leído, podríamos entrar posiblemente a evaluar los procesos implicados en la asociación de grafemas y fonemas, realizando la actividad de lectura de pseudopalabras, y, si encontráramos problemas en esta tarea, podríamos afinar nuestra evaluación realizando una valoración de la conciencia fonológica que tiene nuestro alumno e, igualmente, de los problemas que presenta en la decodificación fonológica y la asociación de fonemas y grafemas; la evaluación efectuada nos hablara, sin duda, de la posible localización de las dificultades de este alumno en estos procesos analizados, otorgándonos también pistas adecuadas para la intervención; también puede ser que no se detecten mayores problemas en estos procesos y se deba continuar la evaluación utilizando actividades que profundizan en otros aspectos, ya sean referentes a procesos psicológicos subyacentes, de integración textual o quizás de autorregulación.

Las flechas discontinuas indican la relación entre la actividad principal (llamémosla así) y aquellas con las que se relaciona y que se efectuarían sobre todo en el caso de constatar la existencia de problemas en los procesos que evalúa la actividad principal. Estas actividades que se encuentran en rectángulos de línea discontinua, sirven, por tanto, al objetivo de permitirnos hilar más fino en nuestra valoración de dónde se encuentran las dificultades que presenta el alumno y, debido a la naturaleza de la evaluación propuesta, también de la intervención que resultará idónea para mejorar el proceso.

Fig. 4. Actividades de evaluación de los distintos procesos implicados en la lectura



7. Procedimiento

Tal y como relatábamos en páginas anteriores, la investigación que ahora presentamos se gestó a medida que se iba construyendo progresivamente el dispositivo de evaluación dinámica EDPL. Nuestro interés por elaborar un procedimiento dinámico de evaluación de la lectura, pensando especialmente en ampliar y optimizar la información que podíamos recabar del alumnado que presentaba especiales dificultades de aprendizaje relacionadas con la lectura, nos había hecho idear y diseñar diversos estudios que pudieran demostrar la validez y la mayor eficacia de la aplicación dinámica de procesos de evaluación en general, y de la evaluación dinámica de procesos implicados en la lectura, en particular, poniendo especial énfasis en la evaluación de los procesos metacognitivos implicados en la comprensión. Uno de estos diseños, que pudimos llevar a la práctica, fue presentado por nosotros en una anterior ocasión (Navarro y Mora, 2002; Navarro y Mora, 2003a), y en él se realizaba una aplicación grupal de un cuestionario sobre metaconocimientos en la comprensión de textos (EML), que originalmente había sido diseñado como parte del dispositivo EDPL y para ser aplicado mediante procedimientos dinámicos de evaluación. Otro de estos estudios que han precedido a la presentación de la actual investigación, fue precisamente el que reflejaba la aplicación dinámica del cuestionario citado anteriormente (Navarro y Mora, 2003b). Este último estudio, que comparte lógicamente más elementos con éste que ahora presentamos, fue realizado con una muestra compuesta por alumnado que presentaba dificultades de aprendizaje de la lectura y, así mismo, en él participaron algunos de los profesores-colaboradores que han intervenido también en el estudio actual.

Paulatinamente, el diseño de nuestra investigación fue adoptando una forma más definitiva, al tiempo que el dispositivo de evaluación dinámica EDPL, ya construido, había sido puesto en práctica mediante la aplicación piloto a la que hacíamos referencia cuando analizábamos el instrumento. Igualmente, las pruebas para evaluar la comprensión de textos y las actitudes relacionadas con la lectura, habían sido elaboradas con la intención de proceder, previamente a su utilización en la fase experimental del estudio, a su aplicación piloto y a una validación por parte de un grupo de expertos. El procedimiento que hemos adoptado, junto con la temporalización aproximada que finalmente hemos podido seguir, lo sintetizamos en el siguiente cuadro. A continuación nos detendremos en cada uno de estos puntos con el fin de desarrollarlos más detalladamente.

Cuadro 32. Procedimiento y temporalización seguidos en nuestra investigación

<i>1. Establecemos contacto con un grupo de profesores/as y orientadores/as para la aplicación piloto de los instrumentos diseñados y para la validación de expertos</i>	septiembre 2003
<i>2. Contactamos con el profesorado y con los orientadores que colaborarían en la fase experimental de la investigación y, a través de estos, con las familias</i>	septiembre/octubre 2003

3. <i>Programamos y comenzamos las sesiones del Grupo de Trabajo (GT) a través del cual se formaría a los profesores aplicadores del dispositivo EDPL y se realizaría un seguimiento y evaluación del proceso de aplicación, así como de los resultados del estudio</i>	septiembre/octubre 2003
4. <i>Programamos y desarrollamos las sesiones de formación para los colaboradores que efectuarían la pasación de las pruebas, tanto a GE como a GC</i>	septiembre/octubre 2003
5. <i>Se comenzó la pasación de estas pruebas en los centros por parte de los profesionales colaboradores</i>	octubre/noviembre 2003
6. <i>Posteriormente se comenzó la aplicación del dispositivo de evaluación dinámica EDPL en el GE por parte de los colaboradores/aplicadores</i>	diciembre 2003
7. <i>Seguimiento y evaluación del proceso de aplicación del dispositivo EDPL por parte del Grupo de Trabajo</i>	diciembre2003/ mayo 2004
8. <i>Aplicación postest de las pruebas</i>	mayo/junio 2004
9. <i>Sesiones de evaluación del Grupo de Trabajo con los aplicadores para valorar el dispositivo de Evaluación Dinámica y la puesta en práctica del mismo</i>	junio 2004
10. <i>Recogida final de datos y comienzo de la determinación de resultados en función de las hipótesis iniciales de trabajo</i>	julio/septiembre 2004
11. <i>Fase final de la investigación. Análisis de datos y elaboración del informe de investigación</i>	octubre 2004/...

1. *Establecemos contacto con un grupo de profesores/as y orientadores/as para la aplicación piloto de los instrumentos diseñados y para la validación de expertos.*

Con los instrumentos de evaluación que fueron elaborados por nosotros, se pretendía operativizar las variables: *nivel de comprensión lectora* y *aspectos de índole socio-emocional*. En este sentido, se trataba de poner en práctica con una muestra piloto una prueba de comprensión de textos (prueba ECO 1, para educación primaria y prueba ECOS, para educación secundaria) y una escala de actitudes, motivación y estrategias de afrontamiento relacionadas con tareas de lectura (escala APSL 1, para el alumnado de primaria y escala APSL, para secundaria). También se realizaría una validación de expertos de los citados instrumentos.

Para la aplicación piloto de las pruebas citadas, se contactó igualmente con el profesorado de los grupos-clase pertenecientes a los dos centros que finalmente han participado en este estudio guía. Se procedió a explicarles en qué consistía la prueba que sus alumnos y alumnas iban a realizar, proporcionándoles el texto introductorio que ofrecía información sobre el contenido de la actividad y sobre

cómo realizarla. Así mismo, se procedió a exponerles someramente los objetivos de nuestra investigación, así como parte del diseño que la estructuraba, ya que se preveía que los grupos-clase participantes en la aplicación piloto podrían participar como un segundo grupo control que debería volver a realizar las pruebas a final de curso, para lo cual necesitaríamos de nuevo la colaboración del profesorado. También se recabó información de los profesores durante estos contactos. El contenido de esta información versó especialmente sobre el alumnado con dificultades de aprendizaje que formaba parte de los grupos, lo que nos era particularmente interesante habida cuenta de que algunos de estos alumnos formarían parte, bien del grupo experimental o bien del grupo control en estos dos centros.

Por otra parte, para la validación de expertos, se hizo entrega de un documento a los profesores y orientadores colaboradores en el que figuraba información acerca del contenido y la estructura de las pruebas a valorar, así como una plantilla de valoración cuantitativa y otra de valoración cualitativa. Cuando los integrantes del grupo de expertos habían realizado sus valoraciones, ya fuera sobre la escala de actitudes o sobre la prueba de comprensión, se procedió a revisar y discutir conjuntamente sus apreciaciones, observaciones, dudas, comentarios o propuestas. Sus valoraciones, junto con los resultados de la aplicación piloto, servirían, como apuntábamos en páginas anteriores, para la mejora de los instrumentos con el fin de su utilización en las fases pretest y postest del estudio con los grupos experimental y control.

2. *Contactamos con el profesorado y con los orientadores que colaborarían en la fase experimental de la investigación y, a través de estos, con las familias.*

Nos pusimos en contacto con los docentes y orientadores que atienden habitualmente al alumnado con dificultades de aprendizaje objeto del estudio, y, a través de estos, con las familias. Explicamos al profesorado los motivos de la investigación y la labor concreta que tendrían que desempeñar en la misma, en función de su pertenencia al grupo de colaboradores/aplicadores del dispositivo de evaluación dinámica, o al grupo de profesores. Este segundo grupo tendría una colaboración mas reducida, que consistiría básicamente en la cumplimentación de un cuestionario sobre expectativas e ideas relacionadas con las dificultades de aprendizaje de la lectura (EIDAL) y en facilitar la pasación de las pruebas pretest y postest. La familia, en principio, sólo tendría que realizar la versión para madres/padres del cuestionario sobre expectativas e ideas con relación a las dificultades de aprendizaje (EIDAL-MP). También se procedería a explicar al alumnado objeto del estudio, de una forma sencilla, la presencia de los profesionales que pasarían las pruebas estandarizadas.

3. *Programamos y comenzamos las sesiones del Grupo de Trabajo (GT) a través del cual se formaría a los profesores aplicadores del dispositivo EDPL y se realizaría un seguimiento y evaluación del proceso de aplicación, así como de los resultados del estudio.*

El GT estuvo formado finalmente por 12 profesores/as y orientadores/as, entre los que se encontraba el coordinador. El inicio de las sesiones de trabajo tuvo como eje principal los objetivos del estudio, así como igualmente los objetivos del GT en

relación a la investigación planteada. Así mismo, se abordaron desde el comienzo algunas cuestiones relativas al contenido y a la metodología propuesta en el dispositivo de evaluación dinámica que debía ser aplicado por los aplicadores. Se repartieron entre los miembros del grupo dos documentos de formación que pretendían facilitar el conocimiento de la fundamentación teórica del dispositivo EDPL, así como la propuesta metodológica que se hacía. El documento de formación 1 abordaba la evaluación dinámica en general y nuestra propuesta de evaluación dinámica de los procesos implicados en la lectura. En el segundo de los documentos se hacía referencia a los procesos contemplados en la propuesta, al esquema de actividades recogido en el dispositivo EDPL y al registro de la información obtenida con su aplicación.

Las primeras cuatro sesiones de trabajo fueron dedicadas a la formación de los aplicadores en aspectos relativos al contenido y la metodología propuesta en el dispositivo. Para ello, una vez repartidos los primeros documentos de formación, se procedió a debatir y analizar su contenido, atendiendo las dudas y observaciones que surgieron entre los aplicadores. Este método de discusión y análisis conjunto fue básico durante el desarrollo de estas sesiones de formación para lograr una cierta unificación en la comprensión y el uso de elementos tan relevantes como los patrones metodológicos de evaluación dinámica, que eran propuestos en el dispositivo para la realización de cada actividad y que recogían la actividad esencial de interacción mediada que debían realizar los aplicadores en la ejecución práctica del dispositivo.

En el transcurso de estas primeras sesiones fue entregado también a los aplicadores un documento de trabajo que contenía el conjunto de actividades de evaluación dinámica que conforman el dispositivo EDPL, así como la propuesta metodológica de mediación y evaluación del proceso de resolución de las actividades y de los procesos metacognitivos implicados en cada actividad. Igualmente se repartieron unos cuadernillos con las *Hojas de registro y evaluación* de cada una de las actividades del dispositivo, así como unas carpetas con las actividades que eran presentadas finalmente a los alumnos y alumnas participantes (hojas para el alumno).

Este material fue objeto de análisis por parte de los miembros del GT, centrandose especialmente la atención en el registro de información durante la aplicación dinámica de las actividades, ya que, aunque en principio el dispositivo EDPL había sido pensado para su aplicación individual, la mayoría de las condiciones reales de aplicación de los integrantes del grupo se caracterizaban por la presencia de varios alumnos a los cuales debía ser aplicado el dispositivo. En este sentido, se analizaron posibles dificultades que podrían surgir en función del número de alumnos y alumnas con los que trabajaban en las aulas de apoyo los integrantes del GT y de las circunstancias concretas del centro donde se ubicaban (número de horas semanales que podían trabajar con sus alumnos, posibilidad de agrupar al alumnado que formaba parte de la muestra, etc.). Muchos de los problemas planteados supusieron un reto para el diseño de investigación realizado, ya que, por una parte, las situaciones prácticas en las que algunos de los miembros del grupo podrían aplicar el dispositivo dificultaban seriamente el trabajo tal y como había sido pensado; por otra parte, suponían la posibilidad de ampliar nuestro análisis de la aplicación a situaciones y contextos con los que en principio no se contaba. Por

ejemplo, aunque EDPL había sido ideado para ser aplicado individualmente, como apuntábamos anteriormente, nuestra experiencia previa y también la realidad del trabajo en sus centros de los miembros del GT, nos confirmaban que la aplicación individual era prácticamente imposible y que, por tanto, muchas de las actividades y propuestas debían ser adaptadas para su realización en pequeño grupo. Esto sin duda enriquecería el análisis que podíamos efectuar del proceso de aplicación, al tiempo que introducía ciertos problemas para evaluar dinámicamente y recoger la información sobre la puesta en práctica de procesos cognitivos y metacognitivos en tareas de comprensión por parte del alumnado.

Posteriormente se comenzaron a analizar en profundidad algunas de las actividades propuestas en el documento de trabajo que recogía la totalidad de las actividades del dispositivo junto con la propuesta metodológica, al tiempo que el coordinador del grupo expuso en estas primeras sesiones una experiencia extraída de la aplicación piloto del instrumento, en la que se abordaba la evaluación de procesos implicados en la asociación grafema-fonema.

4. *Programamos y desarrollamos las sesiones de formación para los colaboradores que efectuarían la pasación de las pruebas, tanto a GE como a GC.*

La búsqueda de colaboradores que tuvieran la cualificación suficiente para poder realizar una aplicación de diversas pruebas, tanto de evaluación de la inteligencia, como de funciones específicas relacionadas con el ámbito escolar, y que, así mismo tuvieran la disposición y la disponibilidad de realizar estas aplicaciones en distintos momentos del curso y en diversos centros educativos de las provincias de Sevilla y Cádiz, no resultó en absoluto tarea fácil. Estos colaboradores fueron seleccionados en el Departamento de Psicología Evolutiva y de la Educación entre aquellos alumnos que ya habían finalizado sus estudios de Magisterio, en la especialidad de Educación Especial, o estaban realizando 5º curso de la licenciatura de Psicopedagogía, y se encontraban de alguna forma ligados al Departamento a través del contacto que mantenían con el profesor Joaquín Mora. Finalmente fueron tres las colaboradoras con las que pudimos contar para la pasación pretest postest de la batería de instrumentos. Dos de estas alumnas habían finalizado sus estudios de Educación Especial en la diplomatura de Magisterio, y una de ellas había finalizado igualmente esta especialidad y cursaba el último curso de la licenciatura de Psicopedagogía.

La necesidad de conocer a fondo las pruebas que serían aplicadas en los centros, al tiempo que la estructura y el contenido de la investigación en la que iban a tomar parte, aconsejó la puesta en marcha de una serie de sesiones de formación (cuatro en total) que tendrían el siguiente contenido:

- Conocimiento y práctica de las pruebas de comprensión de textos ECO 1 y ECOS. Se analizó el contenido de las pruebas y se proporcionó a las colaboradoras el mismo material que fue objeto de valoración por parte del grupo de expertos, con el objetivo de que se familiarizaran con las pruebas y analizaran críticamente el contenido y la estructura de las mismas. También se les propuso que aplicaran la prueba de forma experimental antes del inicio de la aplicación en los centros. Los resultados de esta aplicación fueron analizados

conjuntamente, con el objetivo de mejorar sobre todo aspectos relativos a los tiempos empleados en la realización de los textos de práctica.

- Conocimiento y práctica del test de factor «G» de Cattell. La secuencia seguida fue similar. Las colaboradoras recibieron una copia del manual del test y éste fue leído y analizado individual y conjuntamente. Igualmente, realizaron una aplicación de la prueba previamente a su aplicación en el estudio.
- Conocimiento y práctica de la escala APSL. De igual forma, el contenido y la estructura de esta prueba fue analizada por el grupo de colaboradoras. Se les proporcionó el mismo material con relación a esta prueba que al grupo que realizaría la valoración de expertos. Así mismo, la prueba fue aplicada a personas ajenas a la investigación con el objetivo de familiarizarse con ésta de forma previa a su pasación en los centros.

Uno de los objetivos más importantes de estas sesiones de formación relacionadas con las pruebas, era el de unificar criterios metodológicos, en la medida de lo posible, en el proceso de aplicación. Para ello, se hizo hincapié en el contenido relativo a las instrucciones de aplicación que se hallaban en cada una de las pruebas, así como en algunas premisas básicas que fueron adoptadas con relación a la ayuda que podía ser prestada o a las respuestas que debían ser ofrecidas ante determinadas preguntas.

Además del contenido y la estructura de las pruebas, estas sesiones de formación inicial también fueron dedicadas a conocer someramente el diseño de investigación en el que se insertaban las pruebas. Este conocimiento, sobre todo de algunos aspectos prácticos del estudio, era de vital importancia, ya que a menudo estas profesionales se encontrarían en la necesidad de ofrecer explicaciones a algunos de los profesores que prestaban su colaboración, acerca de los objetivos del estudio o de la necesidad de la aplicación pre y postratamiento para poder realizar las comparaciones oportunas entre los distintos grupos. La experiencia de un estudio anterior nos decía que no siempre es posible que el profesor aplicador, o aquel que colabora en la investigación de una forma más directa, esté presente cuando los colaboradores que van a aplicar las pruebas llegan al centro, por lo que era importante que conocieran bien detalles prácticos referentes al centro o a los grupos-clase participantes. Este conocimiento fue también limitado intencionalmente. Se trataba de que conocieran en profundidad las pruebas que debían aplicar, el esquema de aplicación de las mismas (pretest-postest), la temporalización de la aplicación y los centros a los que debían ir. En cambio, sus conocimientos de los objetivos del estudio fueron limitados, sabiendo solo lo necesario para poder explicar de qué iba el estudio si eran interrogadas en los centros; por otra parte, no conocieron qué sujetos pertenecían a GE o a GC. Tampoco tuvieron información alguna sobre el contenido del dispositivo de evaluación dinámica EDPL, y, así mismo, no tuvieron contacto con el grupo de trabajo en el que los profesores aplicadores se formarían para poner en práctica el tratamiento. De esta forma, se pretendía minimizar el posible riesgo que supondría el conocimiento por su parte de la asignación a GE o GC del alumnado al que pasarían las pruebas, así como, con vistas a la pasación postest, el conocimiento del contenido del tratamiento.

5. *Se comenzó la pasación de estas pruebas en los centros por parte de los profesionales colaboradores.*

Una vez iniciada la aplicación de las pruebas en su fase pretest del estudio, se estipuló necesario tener una sesión semanal de trabajo con las tres colaboradoras mientras durase la pasación de las pruebas. De este modo, durante los aproximadamente dos meses que duró la aplicación, se mantuvo esta reunión, en la que las colaboradoras informaban sobre el proceso de aplicación de las pruebas en los centros, comentaban las dificultades encontradas y proporcionaban las pruebas finalizadas; del mismo modo, se analizaban estas posibles dificultades prácticas con el objetivo de mejorar y unificar el proceso de aplicación, y se planificaba las visitas a los centros para la próxima semana.

6. *Posteriormente se comenzó la aplicación del dispositivo de evaluación dinámica EDPL en el GE por parte de los colaboradores/aplicadores.*

Esta aplicación tuvo una duración media de cuatro meses y medio (unas 18 semanas). Las dificultades a las que antes aludíamos referentes a la necesidad de aplicar el dispositivo EDPL a pequeños grupos de sujetos con dificultades de aprendizaje, o a los horarios y agrupamientos que no siempre facilitaban la labor que nos proponíamos, se vieron extendidas a otros aspectos prácticos de la investigación relativos a la mayor o menor facilidad para aplicar en los centros pruebas, en principio ajenas a las tareas curriculares programadas, a grupos-clase completos. Lógicamente, por cuestiones de ajuste de horarios y agrupamientos, en unos centros fue más fácil que en otros, lo que conllevó en la práctica un desajuste en el ritmo de aplicación de estas pruebas en su fase pretest. Esto a su vez se tradujo en que el comienzo de la aplicación del dispositivo EDPL no se produjo al mismo tiempo para todos los grupos; mientras en algunos centros, las pruebas habían sido ya pasadas a comienzos de diciembre, en otros quedaban algunas por pasar a finales de este mismo mes. Teniendo en cuenta que la aplicación postest podría ocasionar un desajuste parecido, se optó por comenzar la aplicación en los centros que ya habían realizado el pretest, incorporándose el resto a medida que terminaban las pruebas. En la práctica, la variabilidad en el tiempo efectivo de aplicación no ha sido significativa para la mayoría de los centros, aunque posteriormente analizaremos algún caso en particular para el que, por motivos ajenos a los profesores aplicadores, el retraso en la aplicación de las pruebas pretest, así como las dificultades posteriores para llevar a cabo un proceso constante y sistemático de aplicación del dispositivo, puede haber representado un serio *handicap*.

7. *Seguimiento y evaluación del proceso de aplicación del dispositivo EDPL mediante observación sistemática y anotaciones de los aplicadores, así como mediante la visualización y el análisis conjunto de videos en el GT sobre algunas sesiones de aplicación del dispositivo y también mediante sesiones conjuntas de discusión con el responsable del proyecto.*

Después de las primeras reuniones, en las que habíamos tratado de profundizar en los elementos teóricos del dispositivo de evaluación dinámica, así como en algunos aspectos relativos a la metodología propuesta, una vez que los integrantes del GT fueron comenzando la aplicación de EDPL con sus grupos de alumnos, las sesiones

de trabajo estuvieron especialmente dirigidas al análisis de la puesta en práctica del dispositivo en cada uno de los centros, mediante la exposición de los aplicadores, seguida del análisis y discusión conjunta por parte del GT. Esta dinámica seguida fue básicamente la que ha tenido lugar durante la fase de aplicación; es decir, exposiciones de los integrantes del GT sobre el trabajo desarrollado, seguido de observaciones, dudas, comentarios críticos, autoevaluación del proceso, y análisis conjunto por parte de todos los miembros del grupo con el fin de mejorar el proceso.

De la valoración conjunta del grupo de aplicadores se desprende la dificultad que en ocasiones supone el registro de información y la valoración del proceso de resolución de las tareas cuando la aplicación se efectuaba con grupos de alumnos que superaban los diez. Se apuntaron algunas posibles estrategias para recoger el máximo de información posible; entre estas estrategias estaba la utilización del video o de grabadoras, para liberarse de la necesidad de anotar todo lo que ocurría en cada sesión. También se comentó la mayor eficacia que suponía ir anotando ideas o palabras que posteriormente podríamos desarrollar una vez que transcribiéramos a los cuadernillos de cada alumno el desarrollo de la sesión con nuestro análisis sobre la mediación efectuada o la puesta en práctica de habilidades metacognitivas.

También se hizo hincapié durante las exposiciones de los integrantes del grupo, en la importancia de recoger en los cuadernillos que contenían las Hojas de registro y evaluación de cada una de las actividades, sobre todo información referente a los procesos de mediación por parte de los aplicadores, así como a los procesos metacognitivos en la resolución de las actividades por parte del alumnado.

En una sesión posterior se entregó a los aplicadores los cuestionarios EIDAL y EIDAL-MP. Estos cuestionarios, como hemos tenido oportunidad de ver, formaban parte de la investigación que se estaba llevando a cabo y pretendían recoger información acerca de las expectativas e ideas del profesorado y de las familias de los sujetos participantes sobre la atención educativa al alumnado que presentaba dificultades de aprendizaje relacionadas con la lectura y la comprensión. Estos cuestionarios serían cumplimentados especialmente por el profesorado que atendía como tutor/a al alumnado que conformaba los grupos experimental o control del estudio, así como por la familia de estos sujetos. Los colaboradores/aplicadores debían hacer llegar los cuestionarios al profesorado y a las familias. En ocasiones, como alternativa, habría que mantener entrevistas con algunas de las familias con el objeto de facilitar la cumplimentación del cuestionario, ya que se conocía la situación de desventaja socio-cultural —ligada a las dificultades para escribir y leer— en las que se encontraban algunas de ellas.

Se procedió a analizar conjuntamente el contenido de estos cuestionarios para valorar el tipo de información que nos podía ofrecer y su relación con el resto de la información que se estaba obteniendo. Fuimos revisando junto con el GT cada una de las preguntas, tanto del cuestionario preparado para el profesorado como del elaborado para ser cumplimentado por las familias. También se discutieron y analizaron las dificultades provenientes de que en educación secundaria, en ocasiones, el tutor no es el que mejor conoce las dificultades de aprendizaje del alumnado relacionadas con la lectura, por lo que se optó, en función de cada caso

particular, por solicitar la colaboración del profesorado de Lengua o del ámbito Socio-Lingüístico, por ser éstos los que más contacto tenían con el alumnado participante en temas relacionados con la lectura y la comprensión. Se distribuyeron los cuestionarios también entre los profesores participantes del GT que tenían en sus centros algunos alumnos que formaban parte del GC.

En sesiones posteriores se comenzó el visionado de grabaciones de video realizadas por el coordinador del grupo sobre su propia aplicación del dispositivo realizada en uno de los centros participantes. Como ya quedó expuesto en páginas anteriores, la utilización del video en el análisis conjunto de nuestra aplicación de las actividades supone una excelente posibilidad de aprender juntos, tanto de lo que podemos considerar aciertos en nuestra forma de desarrollar el trabajo, como de los errores o aspectos que podríamos mejorar para conseguir los objetivos propuestos. Por otra parte, el video permitía analizar repetidas veces una misma sesión y evaluar el acuerdo entre varios observadores sobre las valoraciones efectuadas en función de los indicadores de evaluación recogidos en cada una de las tareas propuestas. Las sesiones que se expusieron fueron grabadas en el centro Manuel Sánchez Alonso, de Arahal, y en ellas se observaba la aplicación de diversas actividades del dispositivo EDPL. Durante las sesiones de video, se proponía a los integrantes del GT que, observando los indicadores de evaluación de las actividades expuestas, realizaran un análisis sobre la adecuación de la mediación efectuada por el profesor, así como sobre si esta mediación permitía evaluar los indicadores propuestos. El material grabado fue en total de seis horas y media. El trabajo con cada sesión de video en las reuniones del GT oscilaba entre los 20 y los 30 minutos, pasando posteriormente a una puesta en común sobre las observaciones y anotaciones que cada uno había efectuado. Fue especialmente interesante cotejar el proceso de mediación seguido por cada miembro del grupo en la actividad expuesta, con el que se observaba en las distintas grabaciones.

8. *Aplicación postest de las pruebas.* Una vez terminada la fase de tratamiento se volvió a efectuar la pasación de las pruebas contempladas en el diseño por parte de los profesionales que han colaborado a través del Departamento de Psicología Evolutiva y de la Educación.

A medida que se acercaba la fase postest del estudio, en las sesiones del GT se comenzaron a abordar cuestiones prácticas relacionadas con el calendario de visitas a los distintos centros de primaria y secundaria que participaban en el estudio, con el objeto de aplicar las pruebas de evaluación postratamiento. Estas pruebas fueron ya aplicadas al inicio del curso, tanto a los alumnos y alumnas que participaron del GC, como a aquellos que lo hicieron del GE, con los que los integrantes del GT estaban poniendo en práctica el dispositivo de evaluación dinámica de procesos implicados en la lectura. Esta segunda aplicación permitiría poder establecer comparaciones entre el rendimiento de ambos grupos en función del trabajo realizado. Las pruebas serían aplicadas por colaboradores de la investigación ajenos al GT. La colaboración de los integrantes del grupo en esta fase del estudio consistía en facilitar la aplicación de las pruebas en sus respectivos centros.

En cuanto a la aplicación postest de las pruebas, lógicamente nuestra intención fue contar con las mismas profesionales que habían realizado de forma excelente su trabajo en la fase inicial de la fase experimental del estudio; desgraciadamente, tal

y como ya pusimos de manifiesto, sólo una de estas tres colaboradoras podía seguir prestándonos su tiempo y su disponibilidad completa para la aplicación postest. Dado que una persona sola no podía abarcar todo el volumen de trabajo que suponía visitar los trece centros participantes y aplicar tres pruebas distintas en cada centro, debiendo realizar en ocasiones más de tres visitas debido a las ausencias de alumnos o a desajustes en los horarios, se optó por seleccionar otros colaboradores de características similares que pudieran apoyar esta labor. Estos colaboradores deberían igualmente realizar varias sesiones de formación con los mismos objetivos que los planteados con anterioridad para las tres colaboradoras iniciales. Estas sesiones, al igual que las primeras, tuvieron lugar en la facultad de Ciencias de la Educación de la Universidad de Sevilla por espacio de dos semanas, contabilizándose un total de tres. El contenido de estas sesiones fue análogo al expresado para la colaboración inicial. La selección tuvo lugar esta vez entre el alumnado de último curso de Psicopedagogía y el alumnado de la diplomatura de Magisterio. Tres alumnos fueron seleccionados en función de su cualificación y su disponibilidad: un alumno y una alumna de último curso de Psicopedagogía y un alumno de Magisterio. Por tanto, para esta fase postest se pudo contar con cuatro colaboradores.

9. *Sesiones de evaluación del Grupo de Trabajo con los aplicadores para valorar el dispositivo de Evaluación Dinámica y la puesta en práctica del mismo.*

En esta última fase de trabajo con los miembros del GT, la labor desempeñada siguió consistiendo básicamente en el análisis de la aplicación del dispositivo de evaluación dinámica, ya que, este trabajo con el alumnado participante continuaría hasta que fueran aplicadas las pruebas postest en cada uno de los centros. En cualquier caso, la revisión y el análisis conjunto del desarrollo de nuestro trabajo en el estudio permitiría obtener una información muy valiosa sobre las condiciones en las que la aplicación del dispositivo resultó más útil para recabar información sobre los distintos procesos implicados en las dificultades de aprendizaje de la lectura, así como sobre las distintas vías de solución que podrían aplicarse en cada caso. Esta valoración cualitativa sobre la puesta en práctica del dispositivo, y de forma especial sobre los patrones metodológicos de evaluación dinámica que resultaran especialmente idóneos para valorar los procesos metacognitivos implicados en la comprensión; así como el análisis sobre la utilidad y eficacia de EDPL, o las dificultades encontradas para su aplicación, serían objeto de especial atención y valoración en las últimas semanas del GT.

Para facilitar el citado análisis, se entregó a los integrantes del grupo un pequeño cuestionario (Anexo X) que versaba sobre los puntos recogidos en el párrafo anterior, y sobre los que, posteriormente, se discutiría conjuntamente en las sesiones restantes del grupo. También se procedería de igual forma con el documento entregado por el C.E.P. para evaluar el proceso de trabajo y el aprovechamiento del mismo realizado durante el periodo de funcionamiento del GT. El contenido del cuestionario elaborado por nosotros y entregado a los aplicadores era el siguiente:

- *¿Cuáles son las principales dificultades que has encontrado en la aplicación del dispositivo?*
- *¿Qué elementos del dispositivo podrían mejorarse en tu opinión?*

- *¿Qué elementos del dispositivo consideras clave para obtener información valiosa para la posterior intervención?*
- *¿Qué elementos del proceso seguido de evaluación dinámica-intervención consideras clave para explicar la posible mejora en los procesos de lectura?*
- *¿Qué circunstancias han influido o influirían positivamente en una mayor eficacia o utilidad de la aplicación del dispositivo?*

Finalizando el curso, y una vez terminadas las reuniones del GT, se hizo igualmente entrega a los aplicadores de una plantilla (Anexo IX) en la que se solicitaba una valoración del profesorado tutor del alumnado participante acerca del rendimiento escolar en el área del lenguaje escrito al finalizar el curso escolar. Esta valoración debía efectuarse en función de siete criterios establecidos con relación a los formulados desde la Administración educativa y también relacionados con los indicadores valorados en nuestro estudio. Así mismo se solicitaría a los tutores que reseñaran la calificación académica del alumno en el área de Lengua.

10. *Recogida final de datos y comienzo de la determinación de resultados en función de las hipótesis iniciales de trabajo.*

La recogida final de datos se produjo en modo similar a como ya se ha descrito para las pruebas de la fase inicial del estudio, debiendo añadir lógicamente que en esta última fase debía recogerse y registrarse una mayor cantidad de información. Por una parte, las pruebas postratamiento fueron entregadas paulatinamente por los colaboradores que realizaron la aplicación en los centros. Al igual que en la fase pretest, también en esta última fase, durante el periodo que duró la aplicación postest de las pruebas (un mes aproximadamente) tuvimos sesiones semanales con los colaboradores en las que se entregaban por parte de éstos las pruebas finalizadas y se planificaba conjuntamente el calendario de visitas de la semana siguiente. En estas sesiones, se analizaban igualmente posibles dificultades que iban surgiendo — sobre todo, en este último periodo, las dificultades estuvieron relacionadas con las reiteradas ausencias de algunos alumnos de la muestra en las fechas finales del curso, o con circunstancias propias de este periodo como la realización de exámenes—, a las que, en la medida de lo posible, se trataba de dar solución.

Por otra parte, se procedió a recoger los Cuadernos de Hojas de registro y evaluación que habían sido distribuidos entre los integrantes del GT con el objetivo de realizar la valoración, para cada uno de los sujetos del GE, de la aplicación dinámica del dispositivo EDPL mediante las anotaciones sobre la mediación efectuada y el análisis sobre la puesta en práctica de los indicadores de evaluación contemplados. Debido a la dificultad que suponía cumplimentar este cuaderno de hojas de registro y evaluación para cada uno de los alumnos, ya que en ocasiones las anotaciones se habían efectuado en hojas aparte por parte de algunos aplicadores, trasladando posteriormente la información a los cuadernos, éstos fueron entregados a lo largo de los meses subsiguientes. Igualmente espaciada resultó la recogida de las plantillas de valoración del rendimiento en el área de Lengua, que debían cumplimentar los profesores tutores para cada uno de los alumnos participantes. Los motivos en este caso estaban más relacionados con las fechas del curso en las que debía realizarse esta colaboración.

También se procedió a recoger las valoraciones que los miembros del GT realizaron acerca del dispositivo EDPL y de su aplicación experimental. Estas valoraciones, como se recordará, fueron efectuadas a través de un cuestionario de preguntas abiertas que les fue entregado. Aparte de sus valoraciones individuales, plasmadas en este cuestionario, se recogieron valoraciones extraídas del análisis conjunto que se realizó en el GT a raíz de la evaluación del trabajo realizado y de la aplicación del dispositivo en cada una de las circunstancias que rodearon el trabajo desarrollado en los centros.

Por último, se recogieron también de forma paulatina los cuestionarios sobre expectativas e ideas que habían sido cumplimentados por el profesorado y por las familias.

Se ha pretendido adoptar, por tanto, un sistema de *triangulación* o *diversificación* de las fuentes de información obtenida a la hora de evaluar la puesta en práctica experimental del dispositivo EDPL, ya que se recoge información de: (a) los resultados de las pruebas de la batería utilizada; (b) igualmente se recoge información por escrito de los profesionales aplicadores del dispositivo EDPL al cumplimentar las Hojas de registro y evaluación, así como el cuestionario de valoración sobre el dispositivo; (c) el análisis efectuado conjuntamente de las grabaciones de video por parte del grupo de aplicadores en el GT; (d) el contexto de aula donde habitualmente está el alumno evaluado, así como de las circunstancias especiales de algunos alumnos, a través de la información proporcionada por los tutores; (e) las pruebas complementarias, compuestas por los cuestionarios dirigidos a los profesores y a las madres/padres del alumnado; (f) las plantillas de valoración del rendimiento escolar en el área de Lengua que debían cumplimentar los tutores; y, por último, (g) los propios alumnos y alumnas, ya que durante la aplicación del dispositivo EDPL, los aplicadores han recogido información sobre lo que piensa el sujeto acerca de las actividades trabajadas, las dificultades que suponen, lo que se ha aprendido, etc.

11. *Fase final de la investigación. Análisis de datos y elaboración del informe de investigación.*

Varios son los factores que podemos destacar en cuanto al análisis de los datos de nuestro estudio: (a) el número de alumnos y alumnas participantes en el estudio. En este sentido, aunque en el Proyecto de Investigación se recogía la participación aproximada de unos 100 sujetos, la ampliación, en función del interés que su inclusión podía aportar al estudio, a un segundo grupo control compuesto por alumnado que no presentaba dificultades en su aprendizaje, ha supuesto en la práctica que la muestra definitiva haya estado compuesta por 335 estudiantes de educación primaria y secundaria; (b) el diseño de investigación formulado, que requería una pasación inicial de las pruebas y, posteriormente, una aplicación postest de las mismas; (c) el número de pruebas aplicadas en las fases pretest y postest. Han sido aplicadas las tres pruebas que aparecen en el Proyecto en las dos fases a unos 120 estudiantes (aquellos que formaban parte de los grupos experimental y control), dos pruebas a unos 90 estudiantes más y una sola prueba al resto del alumnado participante; (d) la complejidad de la valoración de las pruebas utilizadas. Además del test de factor «G» de Cattell, de fácil corrección, se han utilizado dos pruebas diseñadas por nosotros mismos, una para la evaluación de la

comprensión de textos y otra para la valoración de actitudes y motivación relacionadas con la lectura, cuya complejidad y duración han requerido necesariamente de la elaboración de una matriz de datos igualmente compleja y de grandes dimensiones; (e) la complejidad de la valoración cuantitativa y cualitativa de los protocolos utilizados por el profesorado colaborador en la fase de tratamiento. Se trataba, como apuntábamos anteriormente, de unos Cuadernos de Hojas de registro y evaluación de cada una de las actividades de evaluación dinámica propuestas, donde debía recogerse la valoración del proceso de resolución de las actividades, la mediación efectuada en función de unas pautas previamente proporcionadas, así como la valoración de los procesos metacognitivos puestos en práctica por el alumnado evaluado en el transcurso de las actividades; y (f) la complejidad de la valoración cualitativa de los Cuestionarios de Expectativas e Ideas para padres/madres y para el profesorado del alumnado participante.

En total, han sido analizadas mediante distintos programas del paquete estadístico SPSS, las siguientes pruebas:

Pruebas	(n) fase pretest	(n) fase postest
ECO 1	221	218
ECOS	102	93
APSL	199	192
Factor «G» de Cattell	113	108

- 31 Cuestionarios de expectativas e ideas para el profesorado (EIDAL)
- 29 Cuestionarios de expectativas e ideas para madres y padres del alumnado participante (EIDAL-MP)
- Han sido analizados cuantitativa y cualitativamente 60 Cuadernos de Hojas de Registro y Evaluación, donde por espacio de unas 18 semanas, el profesorado colaborador ha plasmado el proceso de aplicación y los resultados de las actividades de evaluación dinámica propuestas en el estudio

Por otra parte, el análisis de las seis horas y media de grabaciones efectuadas sobre la aplicación del dispositivo de evaluación dinámica, a las que hacíamos referencia en páginas anteriores, no ha podido ser efectuado hasta el momento por psicólogos independientes, tal y como era nuestro deseo, debido a las dificultades prácticas para encontrar personal suficientemente cualificado y con disponibilidad para realizar esta tarea. Ha sido, sin embargo, de gran interés y a todas luces relevante para el desarrollo de las sesiones de evaluación del dispositivo, el análisis conjunto que de estas grabaciones ha efectuado el Grupo de Trabajo formado por el profesorado colaborador del estudio y que estuvo coordinado por el responsable de la investigación. Este análisis de la puesta en práctica del procedimiento propuesto, ha enriquecido de forma significativa el contenido de las reuniones y ha sido útil al efecto de la formación teórica y práctica continuada del Grupo de Trabajo en su aprendizaje y aplicación del dispositivo.

Al igual que ha ocurrido con el análisis de las grabaciones efectuadas en cuanto a las dificultades prácticas para su valoración de forma independiente o ajena al Grupo de Trabajo que realizó la aplicación; la fase que contemplábamos en nuestro proyecto de investigación relativa a la replicación de las pruebas aplicadas pasado un tiempo de la finalización de la aplicación dinámica del dispositivo, no ha podido llevarse finalmente a cabo. Los problemas que han impedido esta actuación tienen que ver

con: (a) el cambio de centro de un número importante de alumnos y alumnas participantes, ya que el alumnado perteneciente a sexto curso de primaria y cuarto curso de secundaria, no era fácilmente localizable ni accesible. También se ha producido una cierta mortandad experimental una vez finalizado el estudio, debido principalmente a motivos relacionados con cambios de centro del alumnado en el transcurso de la etapa o a cambios de profesorado, lo que conllevaba dificultades prácticas para contactar e informar sobre el estudio y sus implicaciones en nuevos centros o a nuevos profesores; (b) el cambio de centro de algunos profesores y profesoras que han colaborado en la aplicación del dispositivo, lo que hacía prácticamente imposible contar con el alumnado participante que había recibido la aplicación del tratamiento; (c) la dificultad —a veces, imposibilidad— de volver a contar con personal colaborador y suficientemente cualificado para realizar las pasaciones de las distintas pruebas en los centros.

En cualquier caso, dado el carácter de la investigación realizada y los objetivos de la misma, es decir, la aplicación experimental de un procedimiento de evaluación dinámica de la lectura con el objetivo de determinar su validez e idoneidad para establecer las condiciones de apoyo y mediación optimizadoras de un mayor desarrollo de la comprensión, así como el volumen de la información obtenida y analizada; la ausencia de una segunda aplicación de las pruebas para evaluar la permanencia de los resultados, aunque hubiera aportado datos de interés, no supondrá en ningún caso un menoscabo de la validez del estudio ni de las conclusiones obtenidas.

Una vez que hemos expuesto secuencialmente el procedimiento seguido en el análisis y evaluación de nuestro estudio, nos disponemos a continuación a mostrar los resultados obtenidos en el mismo. Expondremos esta evaluación del proceso y de los resultados, junto con un primer análisis y valoración de los mismos, en dos capítulos diferenciados. En un primer capítulo se abordarán los resultados y el análisis de los datos obtenidos en las pruebas que conformaban la batería criterio, es decir: las pruebas de comprensión de textos, la escala de ajuste personal-social y la prueba de inteligencia general. En el segundo capítulo afrontaremos la exposición de los datos obtenidos a través de lo que ha constituido nuestra evaluación complementaria; es decir, el análisis y valoración del proceso de aplicación del dispositivo de evaluación dinámica EDPL, así como los resultados obtenidos en el dispositivo. Por otra parte, en este segundo capítulo abordaremos también el análisis diferencial de los valores de predicción del dispositivo dinámico y las pruebas criterio con relación al rendimiento escolar y cognitivo de los sujetos. Por último, expondremos en el último apartado de este segundo capítulo, los resultados obtenidos en los cuestionarios sobre expectativas e ideas de familias y docentes con relación a las dificultades de aprendizaje en la lectura.

RESULTADOS

8. Resultados en las pruebas de comprensión de textos, ajuste personal-social e inteligencia general

8.1 Las pruebas de comprensión de textos

La consideración de las pruebas de comprensión de textos ECO 1 y ECOS como criterio básico en la evaluación de las posibles diferencias en cuanto a rendimiento en lectura entre las fases inicial y final de la investigación para los grupos experimental y control, en función de su participación en la aplicación dinámica del dispositivo EDPL, fue ya explicitada cuando hacíamos alusión a la hipótesis principal de trabajo, así como cuando se describían cada uno de estos instrumentos. En nuestra hipótesis expresábamos de una u otra forma que la evaluación dinámica de procesos implicados en la lectura, especialmente de procesos metacognitivos, así como la aplicación posterior de nuestra propuesta de intervención centrada en los procesos metacognitivos implicados en la comprensión, contribuirían de forma significativa a la mejora del rendimiento en lectura de niños y niñas que presentarían especiales dificultades en su aprendizaje.

Exponíamos igualmente, que la utilización de estas pruebas elaboradas por nosotros mismos, debía servirnos al propósito de establecer comparaciones y, como consecuencia, posibles diferencias entre los resultados obtenidos por los sujetos del GE y del GC. En función de nuestras hipótesis, estas diferencias deberían en buena lógica ser favorables al grupo que participó de la aplicación del dispositivo de evaluación dinámica. En páginas anteriores expusimos algunas de las razones que nos habían llevado a elaborar nuevas pruebas que se mostraran sensibles al trabajo que se pretendía realizar, al tiempo que contemplaran en la mayor medida posible la totalidad de los procesos implicados en la comprensión que habían sido descritos y/o adoptados en nuestra investigación. Al hablar en estos momentos de las pruebas ECO 1 y ECOS como criterio básico para el establecimiento de diferencias pre y postratamiento en el GE, así como de diferencias entre los resultados obtenidos por éste y los mostrados por el GC, debemos hacer referencia también a un aspecto crucial con relación a la validez del estudio; a saber, que el dispositivo de evaluación dinámica EDPL no ha supuesto en medida alguna un entrenamiento para la realización de estas pruebas. Como hemos tenido oportunidad de ver cuando exponíamos el contenido y la metodología del dispositivo aplicado, el trabajo desarrollado con éste está basado especialmente en la interacción mediada que ejerce el evaluador y de la que participa activamente el sujeto; interacción mediada que busca fundamentalmente la apropiación por parte del alumno de patrones de acción cognitiva que optimicen el proceso de comprensión de los textos. Las actividades que realiza el sujeto son esencialmente de tipo reflexivo, en las que debe pensar sobre los procesos que utiliza para comprender o sobre cuáles podría usar para comprender mejor. Igualmente realiza actividades en las que puede poner en práctica habilidades anteriormente modeladas o mostradas por el evaluador, haciendo especial hincapié en el uso de estrategias de control y regulación del proceso de comprensión. Estas actividades, incluso aquellas que externamente guardan similitud con algunas de las recogidas en las pruebas de comprensión, como puede ser la lectura de un texto y el posterior trabajo con cuestiones que aborden diversos aspectos como las

estrategias que podrían ponerse en práctica para comprenderlo mejor o la elección del párrafo que mejor resume el texto, funcionan a modo de «fondo» sobre el que se desarrolla el trabajo de evaluación dinámica de los procesos implicados en la lectura, desarrollándose principalmente en modo oral e interactivo, careciendo por tanto de objetivos como acertar respuestas o aprender reglas para mejorar técnicas específicas, y diferenciándose de este modo, en forma y contenido, de pruebas de opción múltiple de respuesta que son realizadas en modo individual y valoran el producto final del trabajo de comprensión.

Nos disponemos a continuación a mostrar los resultados que hemos hallado en los diversos contrastes realizados. Ya se hizo mención en este sentido a la equivalencia entre los grupos experimental y control en cuanto a su rendimiento inicial en lectura, medido precisamente con estas mismas pruebas. Con relación a la formación de los grupos, esta equivalencia fue calculada teniendo en cuenta la sola designación por parte del profesorado y de los orientadores acerca del alumnado que presentaba dificultades de aprendizaje relacionadas con la lectura.

Esta designación previa de los profesores u orientadores sobre las dificultades de aprendizaje de los sujetos participantes en el estudio, puede ser entendida, y así lo hemos hecho en otras ocasiones (Navarro y Mora, 2003a), como un criterio de validez externa en función de los resultados obtenidos por estos sujetos en comparación con los obtenidos por sus compañeros y compañeras que no presenten dificultades. En páginas anteriores mostramos las diferencias muy significativas que se encontraban entre las puntuaciones del alumnado que participó de nuestra aplicación piloto y las que obtuvo el grupo de alumnos y alumnas con dificultades de aprendizaje. Nosotros hemos tenido en cuenta esta designación del profesorado a la hora de analizar los resultados; sin embargo también hemos querido establecer otros criterios a la hora de determinar los grupos comparables. De esta forma, los resultados obtenidos en las pruebas de comprensión ECO 1 y ECOS, así como en Cattell, nos han servido para efectuar análisis desde una óptica alternativa, quizás más completa o tal vez, como veremos, complementaria. En este sentido, tomaremos hasta tres criterios para analizar los resultados: (a) valoración del profesor/orientador, (b) puntuación en ECO 1 o ECOS, y (c) puntuación en Cattell. Serán las distintas combinaciones entre estos tres criterios, las que nos ofrezcan múltiples posibilidades de análisis que tomen en consideración los distintos elementos que hemos considerado en nuestro estudio. Puede que un alumno cumpla un solo criterio de los que hemos descrito, o dos de ellos, o los tres; es decir, es posible que un alumno haya sido catalogado por el profesorado con dificultades de aprendizaje relacionadas con los procesos de lectura y comprensión, y que sus puntuaciones en ECO 1 o ECOS sean «bajas» e igualmente ocurra con sus puntuaciones en la prueba de inteligencia. Por el contrario, es posible también que otro alumno que ha sido igualmente catalogado con dificultades de aprendizaje, obtenga resultados en la prueba de comprensión que no lo sitúen significativamente por debajo de sus compañeros, e igualmente le ocurra con Cattell. Como puede deducirse, las combinaciones posibles son varias, lo que pensamos puede enriquecer el análisis que nos disponemos a realizar. En cualquier caso, las combinaciones básicas que tendremos en cuenta serán:

(a) Se tendrá en cuenta la designación por parte del profesorado o de los orientadores colaboradores de sujeto con dificultades de aprendizaje de la lectura;

(a) + (b) Se tendrá en cuenta la designación del profesor/orientador y además la obtención de una puntuación «baja» en la prueba de comprensión;

(a) + (b) + (c) Se considerará la designación del profesor/orientador, la obtención de una puntuación baja en la prueba de comprensión y, así mismo, la obtención de una puntuación baja en la prueba de inteligencia general.

Existe, en cambio, un primer problema con el que nos enfrentamos al deber objetivar los criterios que hemos establecido. Por una parte, el factor «G» se distribuye normalmente entre la población, y las puntuaciones obtenidas por los sujetos pueden estimarse e interpretarse en función de las puntuaciones estándar establecidas. De esta forma, dado el carácter de cociente intelectual de desviación que tiene la puntuación final obtenida en la prueba, se puede establecer que una puntuación «baja» podría ser considerada aquella que se encuentra al menos por debajo de una desviación típica de la media, es decir, un CI de 84 o inferior. Se ha descrito así mismo, con relación más específica a la lectura, y a raíz de algunas investigaciones relacionadas con las dificultades en procesos de descodificación y/o de comprensión global de los textos, que algunos sujetos con problemas de lectura no obtienen puntuaciones significativamente más bajas en los test de inteligencia que el resto de sus compañeros sin dificultades (Sánchez, 1999). Nosotros mismos hemos adoptado en cierta medida este criterio como uno de los elementos relevantes a la hora de determinar parte de la muestra del estudio. Estos sujetos suelen ser aquellos catalogados como *disléxicos disfonéticos o fonológicos* y presentan principalmente problemas en los procesos de asociación de grafemas y fonemas. Aunque con la aplicación de nuestras pruebas de comprensión no se pretendía acceder de forma específica a la evaluación de los procesos relacionados con la descodificación, la consideración de la puntuación de corte citada anteriormente podría servirnos de este modo también como criterio de discriminación entre el alumnado que responde de una forma más clara al perfil descrito, en contraposición a aquellos alumnos que presentan dificultades mayormente en la comprensión global de los textos, pudiendo plantear otra serie de problemas relacionados con procesos implicados en la integración textual de los textos, en la integración de estos con experiencias y conocimientos previos, o en la autorregulación de la propia comprensión.

Por otra parte, el nivel de comprensión de textos no está tan «objetivamente» determinado y estandarizado como el factor «G». Sin embargo, las puntuaciones que se han obtenido en la aplicación piloto de las pruebas de comprensión ECO 1 y ECOS pueden servirnos al propósito de establecer estas referencias estándar, máxime cuando los análisis estadísticos expuestos en páginas anteriores revelan la distribución normal de las puntuaciones. De esta forma, debíamos determinar qué puntuaciones serían consideradas «bajas» en función del total de las puntuaciones de la muestra. Como ya se explicó anteriormente, las puntuaciones en las pruebas de comprensión fueron sometidas a una transformación similar a la que caracteriza a las marcas obtenidas en Cattell; en este sentido, las puntuaciones directas de comprensión se transformaron en puntuaciones típicas y éstas a su vez en una escala derivada con media 100 y desviación típica 16. Dada la equivalencia en el proceso de transformación sufrido y la relevancia de la desviación típica en este proceso, nos es válido el argumento y los cálculos efectuados anteriormente con relación al test de Cattell; es decir, considerar una puntuación baja aquella que se encuentre por debajo de una desviación típica de la media. Al igual que en el caso anterior, esta puntuación es de 84 o inferior, lo que viene a coincidir, como puede apreciarse en las tablas 20 y 24, con el percentil 10 de la

muestra de aplicación piloto en el caso de la prueba de comprensión ECO 1, y con un percentil por debajo de 5 para la prueba ECOS.

Una vez establecidos de forma objetiva los criterios para la realización de los distintos análisis tomando en consideración las puntuaciones obtenidas, otro aspecto relevante que hemos de considerar en nuestros contrastes, es el que hace referencia a los distintos subgrupos que consideramos cuando realizábamos la descripción y el análisis de las pruebas, así como a los grupos de aplicación del dispositivo experimental en función de los distintos profesionales que han aplicado EDPL en el grupo de tratamiento. La consideración de estos subgrupos y de las posibles diferencias establecidas en función de los aplicadores, también contribuirán a diversificar, matizar y enriquecer nuestro análisis de resultados.

Nos queda por tanto evaluar las diferencias postratamiento entre las puntuaciones obtenidas, así como la dirección que han tomado estas diferencias y la posible significación estadística de los distintos contrastes. Por otra parte, aunque en apartados posteriores propongamos una reflexión de mayor profundidad en torno a los resultados obtenidos y realicemos a su vez una interpretación de mayor calado de los mismos, abordaremos el sentido y significado que adquieren los datos obtenidos con vistas a la posible verificación de nuestra hipótesis principal, así como de otras consideradas como complementarias.

Resultados en la prueba de comprensión de textos ECO 1 para educación primaria

Comenzaremos por abordar el análisis de la aplicación de la prueba ECO 1 en función del primero de los criterios a los que hacíamos referencia, es decir, la consideración de la valoración de los profesionales que han colaborado en el estudio acerca de las dificultades de aprendizaje del alumnado. Los resultados se muestran en la siguiente tabla.

(a) Criterio de designación previa del profesorado sobre las dificultades de aprendizaje

Tabla 30. Comparaciones pretest-postest para la prueba ECO 1 en función del criterio de designación del profesorado de las dificultades de aprendizaje

*Grupo completo
Prueba t de muestras relacionadas*

Gr.	Dimensiones	N	Pretest	DT	Postest	DT	t	gl	Sign.
GE GC	Puntuación total	30 36	85,198 86,169	8,026 8,907	85,796 92,213	8,849 12,180	-0,363 -2,522	29 35	0,719 0,016*
GE GC	Concepción de la lectura	30 36	92,479 92,479	12,978 11,350	93,270 93,139	11,635 14,076	-0,297 -0,255	29 35	0,769 0,800
GE GC	Metaconocimiento estratégico	30 36	85,733 89,153	9,119 10,543	87,963 93,056	9,287 11,365	-1,204 -1,549	29 35	0,238 0,130
GE GC	Integración textual	30 36	88,408 87,594	9,296 12,926	87,187 95,330	12,874 13,025	0,445 -2,624	29 35	0,660 0,013*
GE GC	Integración texto-conocimientos	30 36	87,939 87,045	11,688 9,128	89,227 92,591	9,287 16,008	-0,555 -2,098	29 35	0,583 0,043*
GE GC	Supervisión y autoevaluación	30 36	90,715 93,025	12,405 12,205	88,926 94,888	10,763 13,896	0,626 -0,797	29 35	0,536 0,431

Podemos observar que mientras que las puntuaciones del GE prácticamente se mantienen, experimentando algunas subidas aunque apenas apreciables, las del GC1 se han incrementado notablemente, mostrando una diferencia de más de seis puntos de promedio en la puntuación total. Esta diferencia, junto con las obtenidas para las dimensiones de *Integración textual* e *Integración texto-conocimientos previos*, llega a ser significativa estadísticamente para un nivel de confianza del 95%. En definitiva, teniendo en cuenta el grupo completo de sujetos que fueron designados como alumnado con dificultades de aprendizaje en la lectura, y que han conformado para este contraste los grupos experimental y control, debemos hacer una primera conclusión relativa a la ausencia de diferencias significativas en el GE entre las aplicaciones pretest y posttest de la prueba, y, por el contrario, constatar que el GC1 sí incrementó significativamente sus puntuaciones entre las dos fases contempladas. Estos resultados pueden ser complementados con un análisis que revele si las diferencias promedio entre las puntuaciones pretest y posttest para ambos grupos resultan estadísticamente significativas. Hemos procedido primeramente a obtener una sola puntuación de cada sujeto, que fuera expresión de la diferencia entre su puntuación inicial y la obtenida en el posttest; posteriormente, mediante la prueba t para muestras independientes y asumiendo la homogeneidad de las varianzas, hemos contrastado las diferencias en las ganancias medias para ambos grupos, obteniendo diferencias favorables al GC que sólo llegan a ser significativas para la dimensión *Integración textual*, con un valor t de 2,190 y una significación asociada de 0,032 para 64 grados de libertad.

Aunque posteriormente ahondaremos en las razones que podrían explicar estos resultados, debemos hacer una importante salvedad con relación a los sujetos que han integrado el GC1 y, como consecuencia, algunos de los diversos contrastes que a continuación iremos exponiendo. En este sentido, cuando hemos hecho referencia al grupo completo, hemos querido destacar especialmente el grupo de sujetos que ha conformado este grupo control, ya que la muestra definitiva que lo ha compuesto finalmente, con una presencia significativamente menor de sujetos que recibían atención educativa especializada en las aulas de apoyo a la integración, así como una cantidad casi significativamente menor de sujetos con deprivación socio-cultural, podría haber introducido sesgos relevantes con relación a los resultados obtenidos.

Habíamos comprobado cuando analizábamos la equivalencia de partida entre los grupos experimental y control, que no existían diferencias significativas en las puntuaciones iniciales en ECO 1. No se encontraron igualmente diferencias significativas con respecto al número total de alumnos que integraban ambos grupos en el contraste inicial. Hemos realizado un nuevo contraste, ahora entre las puntuaciones iniciales del alumnado que recibe AI y aquel que no la recibe para la prueba ECO 1; el hallazgo de diferencias estadísticamente significativas en esta comparación vendría a demostrar que la mayor presencia en número de los sujetos que recibían este tipo de apoyo en el GE se traduciría en desventaja real con respecto al GC1. Los resultados de este contraste nos mostraron que cuando comparábamos las puntuaciones del alumnado que recibía AI con aquellos que recibían apoyo a ciclo o refuerzo educativo (AC), las diferencias llegaban a ser casi significativas: la puntuación media del grupo AI fue de 83,740, con una DT de 9,718; mientras que la media del grupo de refuerzo fue de 88,626, con una DT de 7,159. El contraste con la prueba t para muestras independientes dio un resultado de $t = -1,896$ para 43 gl., con una significación asociada de 0,065. En cambio, cuando se introducían en el contraste los alumnos y alumnas con deprivación socio-cultural, las diferencias,

aunque seguían siendo favorables al grupo de alumnos que no recibía AI, pasaban a ser menores y no se acercaban a la significatividad estadística.

Igualmente podemos recordar que había una presencia mayor de alumnado con privación socio-cultural en el GE; este mayor número de sujetos con desventaja social era casi estadísticamente significativo con respecto al número de alumnos que se encuadraba en el GC1. Hemos realizado un contraste para determinar si esta mayor presencia se traducía para la prueba ECO 1 en diferencias que fueran significativas de partida. Los resultados muestran que cuando comparamos a este grupo con privación con el grupo de alumnos que recibe AC, las diferencias resultan muy significativas ($t = 3,295$; sig. 0,002 para 32 grados de libertad). Incluso cuando la comparación se realiza introduciendo al alumnado que recibía AI, las diferencias siguen siendo estadísticamente significativas ($t = -2,203$; sig. 0,032 para 56 gl.). Estos contrastes muestran la relevancia que podría haber tenido en los resultados la mayor presencia en el GE de alumnado que recibía AI, y de alumnado con privación socio-cultural. Posteriormente, en el capítulo de Discusión retomaremos estos análisis y profundizaremos en las posibles explicaciones de los resultados obtenidos cuando fueron considerados todos los efectivos.

En función de los resultados que hemos expuesto en las líneas anteriores, hemos procedido a realizar un nuevo contraste de las puntuaciones de los grupos experimental y control excluyendo esta vez a los sujetos que no recibían apoyo o que recibían sólo apoyo a ciclo; por tanto, hemos tomado en consideración a los sujetos con dificultades de aprendizaje de la lectura que recibían apoyo a la integración, a los que presentaban discapacidad y a los que habían sido catalogados con privación socio-cultural y manifestaban dificultades en la lectura. La siguiente tabla muestra los resultados obtenidos.

Tabla 31. Comparaciones pretest-postest para la prueba ECO 1 en función del criterio de designación del profesorado de las dificultades de aprendizaje
Grupo reducido
Prueba t de muestras relacionadas

Gr.	Dimensiones	N	Pretest	DT	Postest	DT	t	gl	Sign.
GE GC	Puntuación total	28 15	84,915 81,385	8,220 8,343	84,915 88,412	8,287 11,171	0,000 -1,916	27 14	1,000 0,076
GE GC	Concepción de la lectura	28 15	91,349 89,314	10,921 9,829	93,044 95,645	11,584 15,192	-0,626 -1,468	27 14	0,537 0,164
GE GC	Metaconocimiento estratégico	28 15	85,462 85,733	8,984 9,946	87,135 89,301	9,053 11,826	-0,893 -0,888	27 14	0,380 0,389
GE GC	Integración textual	28 15	88,234 86,454	9,610 13,226	86,925 92,805	12,691 9,780	0,460 -1,549	27 14	0,649 0,144
GE GC	Integración texto-conocimientos	28 15	87,709 79,996	11,559 6,190	88,399 88,583	8,664 15,926	-0,283 -2,000	27 14	0,780 0,065
GE GC	Supervisión y autoevaluación	28 15	91,002 89,820	12,676 11,855	87,648 91,609	9,951 13,287	1,192 -0,521	27 14	0,244 0,610

Cuando se han realizado nuevos contrastes tomando en consideración sólo al alumnado con DA o privación que recibían atención educativa especializada en las aulas de

apoyo a la integración, o al alumnado que presentaba discapacidad; podemos observar que el número de efectivos del GC1 ha descendido notablemente cuando se han eliminado para este contraste los sujetos que no presentaban discapacidad o privación social, o que así mismo no recibían atención educativa especial. En cambio, el número de sujetos del GE solo se ha reducido en dos. Los resultados no han sufrido por tanto mucha variación para este grupo; en cambio, el GC1 no obtiene ahora diferencias que resulten estadísticamente significativas, aunque mantiene unas diferencias importantes favorables al postest, tanto en la puntuación total como en algunas de las cinco dimensiones contempladas.

Contrastes en función de los distintos subgrupos considerados en la muestra

Nos proponemos ahora analizar los resultados obtenidos en la prueba de comprensión ECO 1 teniendo en cuenta los diversos subgrupos que hemos tomado en consideración en anteriores análisis. En este sentido, dado que la comparación entre sujetos que reciben AI y aquellos que no asisten a las aulas de apoyo, parece de alguna forma sesgada, y que, por otra parte, en el GE hay una presencia significativamente mayor de sujetos que se encuentran en la primera de estas condiciones, nos resulta especialmente relevante la realización de un contraste entre las puntuaciones pretest y postest de aquellos alumnos que reciben AI en los grupos experimental y control. Los resultados de esta comparación, junto a los obtenidos para el resto de los subgrupos contemplados, se muestran en la siguiente tabla.

Tabla 32. Comparaciones pretest-postest en la prueba ECO 1 (puntuación total) para los distintos subgrupos contemplados en la muestra (DA = dificultades de aprendizaje; DIS = discapacidad; AI = recibe apoyo a la integración; DEP = privación socio-cultural; AC = recibe apoyo a ciclo o refuerzo educativo; AF = agrupamientos flexibles; NA. = no recibe apoyo)

Prueba t para muestras relacionadas

Grupos	N	Pretest	DT	Postest	DT	valor t	gl	significación
GE DA	11	86,917	8,569	86,713	11,887	0,082	10	0,936
GC1 DA	22	88,752	8,551	94,053	12,774	-1,721	21	0,100
GE DIS	7	84,354	6,781	85,635	6,456	-0,487	6	0,643
GC1 DIS	4	80,749	9,041	82,431	11,139	-0,345	3	0,753
GE AI	19	85,737	8,762	85,973	8,950	-0,124	18	0,903
GC1 AI	8	80,469	7,717	81,590	8,722	-0,333	7	0,749
GE DEP	12	84,114	8,560	85,048	7,335	-0,289	11	0,778
GC1 DEP	10	82,656	8,2204	92,076	10,198	-1,839	9	0,099
GE NA, AC, AF	11	84,266	6,863	85,490	9,098	-0,385	10	0,708
GC NA, AC, AF	28	87,798	8,661	95,248	11,385	-2,571	27	0,016*

Podemos observar que los contrastes realizados revelan la presencia de significación estadística para el GC1 sólo cuando la comparación pretest-postest se efectúa con el grupo de sujetos formado por aquellos que no reciben apoyo a la integración (n = 10), que reciben apoyo a ciclo (n = 15), o que forman parte de agrupamientos flexibles (n =

3). Así mismo, las diferencias son mayores, aún sin llegar a ser significativas, cuando los contrastes son efectuados con el grupo de sujetos con dificultades de aprendizaje (DA), del que, como apuntábamos anteriormente, se han extraído aquellos alumnos que presentaban también discapacidad o deprivación socio-cultural, y con el grupo de sujetos con deprivación (DEP). En cambio, para el alumnado con discapacidad y para el grupo que recibe AI, no se han encontrado diferencias reseñables en el GC1. Por lo que respecta al GE, las diferencias obtenidas entre las puntuaciones pretest y postest para cada uno de los subgrupos contemplados, confirman los resultados obtenidos para el contraste realizado con el grupo completo. Se observan pequeños incrementos en las puntuaciones postest para casi todos los subgrupos analizados, que en ningún caso se acercan a la significatividad estadística.

Podemos constatar también mediante la exploración de los datos que el número de alumnos y alumnas que sólo fue catalogado con dificultades de aprendizaje por sus profesores, es decir, en ausencia de discapacidad o deprivación socio-cultural, es muy superior en el GC, lo que, en contra de lo que podría suponerse, dadas las características de la mayoría de estos sujetos, es decir, alumnos y alumnas que presentan dificultades relacionadas con los procesos de comprensión y que reciben apoyo a ciclo o refuerzo educativo, pero que, en principio, sus dificultades no conllevan la necesidad de una atención educativa específica en las aulas de apoyo a la integración, podría haber contribuido al incremento de las puntuaciones postest obtenidas por el GC1. Por otra parte, podemos observar igualmente que el número de sujetos que reciben AI en el GE es notablemente superior al del GC1. Ya habíamos analizado anteriormente la posible significación estadística de la mayor presencia en el GE de este subgrupo. Los resultados obtenidos con la prueba chi-cuadrado revelaban la existencia de diferencias significativas. Así mismo, hemos realizado este análisis para evaluar la mayor presencia en el GC de sujetos que reciben apoyo a ciclo. Los resultados muestran un valor para chi-cuadrado de 8,222; con una significación estadística de 0,042 para 3 gl.

Cuando realizábamos los contrastes para evaluar la equivalencia de los grupos experimental y control en educación primaria con relación al sexo, apuntábamos que, aunque la comparación entre ambos grupos podía ser legítima, a tenor de que no se hallaban diferencias estadísticamente significativas para la mayor presencia en el GE de niños con respecto al GC; existía la posibilidad, dadas estas diferencias en cuanto al número de niños que conformaban cada uno de los grupos, de que los resultados estuvieran en parte influidos por esta circunstancia. Recordemos que en el GE había 11 niñas y 19 niños que realizaron la prueba, mientras que en el GC1 había 17 niñas e igual número de niños que en el GE. Hemos realizado para la prueba ECO 1 un contraste de las puntuaciones pretest-postest obtenidas por chicos y chicas, tomando posteriormente en consideración este hecho en nuestros análisis.

Podemos observar que el único contraste que resulta estadísticamente significativo es el que tiene lugar entre las puntuaciones pretest y postest que obtuvieron las niñas del GC1. Concretamente, la diferencia observada para este subgrupo parece explicar la mayor parte de la que quedó establecida para el GC1 en su conjunto. Hemos querido, no obstante ahondar un poco más en los datos y saber qué proporción del subgrupo de niñas del GC1 estaba formado por alumnas que recibían apoyo a la integración, ya que parece ésta una variable determinante en los resultados obtenidos. Nuestra exploración de los datos nos revela que sólo 4 alumnas, el 23,53%, recibía AI; el análisis de las puntuaciones pretest y postest de estas cuatro alumnas nos muestra unos resultados bien

Tabla 33. Comparaciones pretest-postest por sexos en la prueba ECO 1 (puntuación total)

Prueba t para muestras relacionadas

Grupos		N	Pretest	DT	Postest	DT	valor t	gl	significación
GE	Niñas	11	84,674	7,233	84,063	7,377	0,215	10	0,834
	Niños	19	85,501	8,628	86,799	9,646	-0,630	18	0,536
GC1	Niñas	17	86,257	9,800	96,020	14,781	2,214	16	0,042*
	Niños	19	86,091	8,298	88,806	8,266	-1,299	18	0,210

distintos: pretest = 84,114; postest = 82,992. Por tanto, dentro del grupo de niñas del GC1, parece ser que el subgrupo formado por aquellas que no recibían ningún tipo de apoyo o recibían AC o AF, ha sido el que explicaría las notables diferencias halladas. Por el contrario, las puntuaciones del GE vuelven a no dejar entrever ningún atisbo de incremento significativo; es más, en esta ocasión, las puntuaciones decrecen para las niñas. Para el grupo de niños experimentan una ligera subida, aunque sin llegar a acercarse a la significación estadística. Cuando hemos realizado el análisis anterior para comprobar cómo estaba formado el grupo de niñas, los resultados nos muestran que 5 de las 11 niñas, el 55%, recibía AI, con unas puntuaciones que también experimentan decremento: pretest = 86,469; postest = 84,674. La gran diferencia con respecto al GC1, además de tener una proporción mayor de niñas que reciben AI y, por tanto, con mayores dificultades, es que el resto de niñas, es decir, aquellas que reciben AC o AF no obtiene ganancias entre sus puntuaciones pretest y postest, manteniéndolas prácticamente sin cambios.

Contrastes en función de los grupos de aplicación del dispositivo

Nos disponemos a continuación a realizar los contrastes en función de los distintos aplicadores que han intervenido en el estudio. Lógicamente, en ningún caso se tratará de comparar la calidad del trabajo realizado ni la pericia como docente; se trata en este caso de aproximarnos a un análisis cuantitativo —siempre limitado en nuestra opinión— de las diferentes formas de aplicación experimental del dispositivo. Como ya se apuntó en páginas anteriores, durante las sesiones de formación de los aplicadores se había tratado de poner de relieve la importancia de aunar criterios en cuanto al seguimiento de los patrones metodológicos de evaluación dinámica contenidos en el manual de EDPL, así como de seguir el itinerario marcado en cuanto a las actividades a realizar. Con todo, y a pesar de que en las posteriores sesiones del Grupo de Trabajo, donde se realizaba un seguimiento y evaluación del proceso de aplicación del dispositivo, se volvía reiteradamente sobre la importancia de realizar una aplicación lo más ceñida posible a la propuesta para el estudio, es inevitable que en la práctica de la aplicación de cualquier programa educativo, salga a relucir un componente personal y relativo al estilo docente y de relación interpersonal. Las diferencias que este componente personal del docente pueda introducir en la práctica de una intervención educativa como la nuestra, nos parecían lógicas al tiempo que ineludibles, por lo que sólo podíamos ejercer un cierto control antes y durante la aplicación (tratando, en la medida de lo posible, que la aplicación siguiera unas pautas comunes en cuanto a

criterios metodológicos e itinerario de acción), así como posteriormente, tratando de conocer la posible influencia que los distintos aplicadores han podido imprimir en los resultados obtenidos. En la siguiente tabla mostramos las comparaciones por grupos a los que un mismo evaluador aplicó el dispositivo en el grupo experimental.

Tabla 34. Comparaciones pretest-postest para la prueba ECO 1 por grupos de aplicadores Estadísticos de muestras relacionadas

Grupos	Dimensiones		Media	N	DT	t	gl	Sign.
Grupo A CEIP Andalucía	Puntuación total	Pretest	83,1794	9	7,0925	0,149	8	0,885
		Postest	82,6810		6,5816			
	Concepción sobre la lectura	Pretest	92,4796	9	11,8701	1,512	8	0,169
		Postest	87,2040		7,9134			
	Metaconocimiento estratégico	Pretest	82,4616	9	8,2708	-0,206	8	0,842
		Postest	83,2051		6,7837			
	Integración textual	Pretest	85,9656	9	9,6951	0,450	8	0,665
		Postest	84,3370		8,5502			
	Integración texto-conocimientos	Pretest	89,7288	9	10,0701	-0,164	8	0,873
		Postest	90,4445		8,1754			
	Supervisión y Autoevaluación	Pretest	85,9454	9	7,0705	0,000	8	1,000
		Postest	85,9454		7,0705			

Grupos	Dimensiones		Media	N	DT	t	gl	Sign.
Grupo B CEIP Antonio Machado	Puntuación total	Pretest	85,796	6	6,142	0,000	5	1,000
		Postest	85,796		7,057			
	Concepción sobre la lectura	Pretest	92,479	6	12,259	No se puede calcular t		
		Postest	92,479		12,259			
	Metaconocimiento estratégico	Pretest	83,948	6	3,455	-0,791	5	0,465
		Postest	86,179		7,330			
	Integración textual	Pretest	88,408	6	5,983	-0,277	5	0,793
		Postest	89,630		12,623			
	Integración texto-conocimientos	Pretest	87,939	6	9,768	0,210	5	0,842
		Postest	86,866		6,652			
	Supervisión y Autoevaluación	Pretest	97,870	6	16,246	0,439	5	0,679
		Postest	93,398		11,999			

Grupos	Dimensiones		Media	N	DT	t	gl	Sign.
Grupo C CEIP El Ruedo	Puntuación total	Pretest	85,422	3	4,668	1,512	2	0,270
		Postest	79,441		5,644			
	Concepción sobre la lectura	Pretest	92,479	3	13,706	0,000	2	1,000
		Postest	92,479		13,706			
	Metaconocimiento estratégico	Pretest	81,718	3	7,726	-0,378	2	0,742
		Postest	83,948		3,863			
	Integración textual	Pretest	98,180	3	7,328	5,196	2	0,035*
		Postest	76,193		7,328			
	Integración texto-conocimientos	Pretest	86,866	3	7,437	1,000	2	0,423
		Postest	82,572		7,437			
	Supervisión y Autoevaluación	Pretest	84,454	3	7,745	-2,000	2	0,184
		Postest	93,398		13,415			

Tabla 34. Comparaciones pretest-postest para la prueba ECO 1 por grupos de aplicadores (Continuación) Estadísticos de muestras relacionadas

Grupos	Dimensiones		Media	N	DT	t	gl	Sign.
Grupo D CEIP El Ruedo	Puntuación total	Pretest	90,281	2	7,929	No se puede calcular t		
		Postest	81,310		7,929			
	Concepción sobre la lectura	Pretest	96,436	2	16,786	0,000	1	1,000
		Postest	96,436		16,786			
	Metaconocimiento estratégico	Pretest	92,870	2	9,463	3,000	1	0,205
		Postest	82,833		4,731			
	Integración textual	Pretest	98,180	2	0,000	1,667	1	0,344
		Postest	79,858		15,546			
	Integración texto-conocimientos	Pretest	75,057	2	4,554	-3,000	1	0,205
		Postest	84,719		0,000			
	Supervisión y Autoevaluación	Pretest	113,521	2	9,486	3,000	1	0,205
		Postest	93,398		18,972			

Grupos	Dimensiones		Media	N	DT	t	gl	Sign.
Grupo E CEIP Lope de Rueda	Puntuación total	Pretest	95,888	3	11,868	-1,192	2	0,355
		Postest	102,617		10,278			
	Concepción sobre la lectura	Pretest	100,393	3	27,412	0,000	2	1,000
		Postest	100,393		13,706			
	Metaconocimiento estratégico	Pretest	95,101	3	13,929	-1,890	2	0,199
		Postest	106,253		11,589			
	Integración textual	Pretest	95,737	3	8,462	-0,277	2	0,808
		Postest	98,180		19,390			
	Integración texto-conocimientos	Pretest	99,748	3	19,677	-0,756	2	0,529
		Postest	104,041		11,155			
	Supervisión y Autoevaluación	Pretest	93,398	3	13,415	-0,277	2	0,808
		Postest	97,870		15,490			

Grupos	Dimensiones		Media	N	DT	t	gl	Sign.
Grupo F CEIP Sánchez Alonso	Puntuación total	Pretest	81,150	7	7,534	-1,662	6	0,148
		Postest	86,597		6,267			
	Concepción sobre la lectura	Pretest	87,957	7	8,973	-2,121	6	0,078
		Postest	98,132		12,689			
	Metaconocimiento estratégico	Pretest	87,135	7	10,529	-0,880	6	0,413
		Postest	90,958		5,058			
	Integración textual	Pretest	81,428	7	6,971	-1,591	6	0,163
		Postest	90,851		14,657			
	Integración texto-conocimientos	Pretest	84,719	7	12,333	-0,411	6	0,695
		Postest	87,479		8,999			
	Supervisión y Autoevaluación	Pretest	85,732	7	7,170	1,000	6	0,356
		Postest	81,899		5,070			

La inspección de los datos de aplicación del dispositivo en función de los distintos profesionales que han participado en su puesta en práctica experimental, pone de manifiesto la existencia de diferencias notables entre las puntuaciones obtenidas por los grupos. El factor *aplicador* podría de esta forma haber tenido una cierta influencia en las puntuaciones globales del grupo experimental. Estas diferencias se tornan especialmente relevantes entre, por una parte, los grupos de aplicación C y D, donde se han producido las puntuaciones significativamente más bajas del conjunto y, por otra, los grupos E y F de la muestra donde se han obtenido las diferencias más importantes a favor del tratamiento. Estas diferencias han tenido lugar para estos dos grupos en la práctica totalidad de las dimensiones. Por otra parte, en los otros dos grupos de

aplicación, como puede observarse, las puntuaciones de los grupos no han experimentado prácticamente cambios.

Hemos querido comprobar estadísticamente la relevancia de las diferencias observadas. En este sentido, hemos realizado un contraste entre las diferencias de medias para cada uno de los grupos de aplicación. La comparación, tomando en consideración los seis grupos de forma conjunta, arroja un valor F de 1,645 con una significación estadística asociada de 0,186 para 5 grados de libertad. En cambio, cuando hemos calculado los estadísticos de contrastes *post hoc* para comparaciones múltiples por pares de grupos, mediante la utilización del estadístico DMS, que utiliza pruebas t para llevar a cabo todas las comparaciones por pares entre las medias de los grupos, aunque sin efectuar ninguna corrección de la tasa de error por el hecho de realizar múltiples comparaciones, hemos hallado diferencias significativas entre los grupos D y F, obteniendo un valor de diferencia de 14,418, con una significación estadística de 0,046. Así mismo, la diferencia entre los grupos D y E llega a ser casi significativa (sig. 0,056), como la que hay entre los grupos C y F (sig. 0,065). Esta significatividad de las diferencias no se ha encontrado cuando se han utilizado otros estadísticos que sí efectúan correcciones de la tasa de error en función de las múltiples comparaciones (Scheffé: el valor de la diferencia entre los grupos D y F fue de 14,418; sig. 0,508). En este sentido, aunque las diferencias encontradas parecen importantes, el número de sujetos que componen los grupos parece influir decisivamente para que estas diferencias no tengan un reflejo más claro en los contrastes estadísticos.

Resulta de interés para nosotros, por razones relativas a las condiciones de aplicación en las que profundizaremos más adelante, realizar un nuevo contraste prescindiendo de los grupos A, C y D. En esta comparación, obviamente, los resultados son más positivos y favorables a la hipótesis de partida, aun sin llegar a ser estadísticamente significativos. Para 16 sujetos, hemos obtenido unas puntuaciones medias de 85,655 en el pretest y de 89,300 en el postest; el valor t observado ha sido de -1,790, con una significación asociada de 0,094 para 15 gl.

Por lo que respecta al nivel educativo, recordamos que los análisis de equivalencia realizados con anterioridad reflejaban la ausencia de diferencias reseñables entre los grupos experimental y control para la comparación con el grupo completo de sujetos. Hemos obtenido los análisis para las posibles diferencias entre las fases inicial y final del estudio en función de esta variable. Tal y como hicimos en los análisis de equivalencia, los contrastes los hemos efectuado por ciclos. Los resultados para el GE en el segundo ciclo ($n = 5$) no muestran diferencias estadísticamente significativas para la puntuación total, con una media pretest: 87,814 (DT = 4,074) y una postest de 86,020 (DT = 8,915), con un valor t para muestras relacionadas de 0,547 y una significación de 0,614 para 4 gl. Para el tercer ciclo ($n = 25$), tampoco se obtuvieron diferencias significativas; la media inicial del GE fue de 84,674 (DT = 8,565) y la postest de 85,751 (DT = 9,020), con un valor t de -0,574 y una significación de 0,572 para 24 gl. Por su parte, el GC no muestra diferencias significativas para el segundo ciclo de primaria ($n = 12$), obteniendo una media pretest en la puntuación total de 85,422 (DT = 9,039) y una postest de 87,665 (DT = 8,133), con un valor t de -0,793 y una significación de 0,445 para 11 gl.; mientras que para el tercer ciclo ($n = 24$) sí se reflejan diferencias estadísticamente significativas, tanto en la puntuación total, con un pretest de 86,543 (DT = 9,012) y un postest de 94,487 (DT = 13,339), $t = -2,427$; sig. 0,023 para 23 gl.,

como en dos de las dimensiones contempladas: *Integración textual* e *Integración texto-conocimientos*.

(b) *Criterio del profesorado + puntuación inicial en la prueba ECO 1*

Nos disponemos ahora a analizar los resultados en función de este segundo criterio que, como decíamos, contempla, además de la valoración previa del profesorado o de los orientadores acerca de la condición de dificultad de los sujetos, la puntuación obtenida por éstos en la fase inicial del estudio en la prueba (puntuación total). Esta puntuación debía igualar o estar por debajo de 84 para ser considerada en estos contrastes. Los resultados para los grupos experimental y control los exponemos en la siguiente tabla.

Tabla 35. Comparación pretest-postest para la prueba ECO 1 en función del criterio de designación del profesorado más la puntuación inicial en la prueba (puntuación ≤ 84)

Prueba t para muestras relacionadas

Gr.	Dimensiones	N	Pretest	DT	Postest	DT	T	Gl	Sign.
GE GC	Puntuación total	11 13	77,742 76,910	4,646 4,449	83,859 90,885	4,626 12,083	-2,382 -3,991	10 12	0,038* 0,002**
GE GC	Concepción de la lectura	11 13	88,882 88,218	9,603 8,915	93,199 91,870	11,977 14,966	-1,491 -0,693	10 12	0,167 0,502
GE GC	Metaconocimiento estratégico	11 13	78,879 81,031	6,315 4,851	86,179 91,841	7,330 12,169	-2,292 -2,880	10 12	0,045* 0,014*
GE GC	Integración textual	11 13	84,188 78,448	11,092 8,130	86,853 93,670	8,893 10,594	-0,599 -3,959	10 12	0,563 0,002**
GE GC	Integración texto-conocimientos	11 13	80,620 79,764	8,284 5,359	86,475 93,142	7,108 15,193	-1,614 -3,171	10 12	0,138 0,008**
GE GC	Supervisión y autoevaluación	11 13	84,861 88,238	6,768 10,302	87,300 92,366	9,223 10,189	-0,803 -1,000	10 12	0,441 0,337

Como podemos observar, cuando hemos adoptado el criterio de la puntuación obtenida en ECO 1 en la fase inicial del estudio, unido a la valoración previa del profesorado, obtenemos diferencias estadísticamente significativas a favor del GE para la puntuación total y también para la dimensión *Metaconocimiento estratégico*. Además, el contraste muestra que las puntuaciones postest son claramente superiores a las iniciales en la práctica totalidad de dimensiones evaluadas. Podríamos, por tanto, argumentar en primera instancia que el dispositivo de evaluación dinámica aplicado ha tenido un efecto significativo a favor de la hipótesis formulada cuando la aplicación ha ido dirigida precisamente al alumnado con mayores dificultades de lectura y comprensión. Sin duda, este dato es favorable y será de interés profundizar con posterioridad en su análisis; sin embargo, la inspección de los datos nos ofrece una información rotunda; a saber, las diferencias entre las puntuaciones medias obtenidas por el grupo control, en idénticas condiciones en cuanto al criterio seguido, resultan ser aún de mayor significatividad estadística. Las diferencias para el GC resultan significativas para la puntuación total y para las dimensiones *Metaconocimiento estratégico*, *Integración textual* e *Integración texto-conocimientos previos*.

Hemos contrastado las diferencias medias para evaluar la posible significatividad de las ganancias obtenidas por uno de los grupos. Observando la magnitud de las diferencias,

podemos observar que éstas favorecen generalmente al GC, lo cual ha sido confirmado en parte mediante el análisis realizado con la prueba t para muestras independientes, obteniendo tan sólo para una de las dimensiones contempladas, *Integración textual*, diferencias estadísticamente significativas con relación a las ganancias medias (valor de $t = 2,147$; sig.0,043 para 22 gl.).

Por otra parte, resulta de interés conocer quien compone los grupos que acabamos de contrastar; sabemos que se trata de alumnos y alumnas que fueron catalogados por sus profesores, o por los orientadores que han colaborado en la investigación, como sujetos con dificultades de aprendizaje en la lectura, y que, así mismo, obtuvieron una baja puntuación inicial en la prueba de comprensión que nos ha servido para evaluar el rendimiento en lectura. Hemos querido conocer además cómo estaban distribuidos los efectivos de estos grupos en función de las categorías que fueron contempladas en anteriores análisis, es decir: dificultades de aprendizaje, deprivación socio-cultural o discapacidad, con relación al tipo y grado de sus problemas de aprendizaje; y si recibe apoyo a la integración, apoyo a ciclo/refuerzo educativo, o participa de agrupamientos flexibles, en cuanto al tipo de ayuda educativa que se le ofrece.

La exploración del grupo experimental nos muestra que 5 alumnos responden al perfil de deprivación socio-cultural, participando cuatro de ellos de agrupamientos flexibles y recibiendo apoyo a la integración el otro; 3 alumnos presentaban discapacidad —dos de ellos con deprivación socio-cultural—, recibiendo los tres apoyo a la integración; y 3 alumnos fueron catalogados sólo con dificultades de aprendizaje, recibiendo apoyo a la integración. Por su parte, los trece alumnos del GC se distribuyen de la siguiente forma: 5 alumnos presentaban deprivación social, de los cuales 3 participaban de agrupamientos flexibles y 2 recibían apoyo a la integración; 3 alumnos presentaban discapacidad —uno de ellos con deprivación—, recibiendo los tres apoyo a la integración; y 5 alumnos fueron catalogados con dificultades de aprendizaje, de los cuales 1 recibía apoyo a la integración, otro apoyo a ciclo y tres no recibían ningún tipo de apoyo. En definitiva, la consideración de los integrantes de estos dos grupos en función de la tipología de sus dificultades o de las condiciones de apoyo, no parece establecer unas diferencias claras entre ambas formaciones; el número de alumnos con discapacidad es igual en ambos grupos y el de alumnos que reciben apoyo a la integración es similar (7 en el GE y 6 en el CG); así mismo, el número de alumnos con deprivación social también guarda esta relación entre los dos grupos; por otra parte, sí habría que considerar que en el GC hay 3 alumnos que no reciben apoyo y 1 que recibe apoyo a ciclo (recuérdese que esto venía a significar en nuestro estudio que este alumnado no precisaba una atención educativa tan específica, de lo que se deduce a tenor de las investigaciones anteriormente aludidas, que su mejora postest habría requerido igualmente una intervención menos específica). Hemos contrastado las puntuaciones pretest-postest de los grupos excluyendo a los sujetos que no precisaran apoyo específico o que tan sólo recibiera apoyo a ciclo. Los resultados muestran que para el GE no cambia nada, pues la totalidad de sus efectivos cumple las condiciones de recibir AI, presentar discapacidad o deprivación socio-cultural, además de dificultades de aprendizaje relacionadas con la comprensión; por su parte, el GC1, al igual que pasara con el grupo completo, ve reducido sus efectivos, ya que cuatro alumnos no cumplían las anteriores condiciones. El nuevo grupo, formado por 9 sujetos obtiene para la puntuación total una puntuación pretest de 75,703 (DT = 4,196) y una postest de 88,911 (DT = 13,846), con un valor t de -2,668 y una significación asociada de 0,028 para 8 grados de libertad. Por tanto, la significatividad del contraste sigue estando

patente, aunque ahora se produzca con un nivel de confianza menor. En cuanto a las dimensiones, ahora los resultados muestran la ausencia de diferencias significativas para la dimensión *Metaconocimiento estratégico* y la significación a un nivel 0,05 de *Integración textual* e *Integración texto-conocimientos*.

Hemos visto que tanto en el GE como en el GC, el grupo de alumnos que puntuaba inicialmente por debajo de una desviación típica sobre la media, y que, por tanto, cumplía los dos primeros criterios a los que aludíamos al comienzo de este apartado, obtenía resultados significativamente mejores en su realización posttest de la prueba; sin embargo no sabemos aún cómo se comportaron sus compañeros que obtuvieron puntuaciones más altas de partida. La lógica parece indicar que sus resultados tendrían que ser mayores en el posttest a los obtenidos por los alumnos que en principio manifestaban mayores dificultades —exceptuando el caso de que la prueba utilizada causara un efecto «techo», ya fuera por su dificultad o su facilidad, por el que el alumnado no pudiera superar de forma significativa puntuaciones bajas o altas conseguidas en la fase inicial—. Parece en principio que la amplitud de la prueba en cuanto al grado de posible margen para la mejora, no hace muy viable el hecho de no poder incrementar las puntuaciones iniciales, aunque sí podría haber ocurrido que las puntuaciones de nivel medio obtenidas por algunos sujetos con dificultades no hayan podido ser superadas por éstos en función de un cierto efecto techo de distinto tipo al que describíamos anteriormente. Este efecto habría actuado, por una parte, en relación al margen de mejora posible que pudieran tener estos sujetos en el periodo de duración del estudio, y, por otra —para el GE—, en relación a la validez del tratamiento aplicado para suscitar cambios significativos en ese grupo de alumnos. En este caso al que aludimos, la aplicación del dispositivo EDPL no habría causado en estos alumnos una mejora significativa de su competencia lectora. Hemos analizado, tanto para el GE como para el GC el comportamiento de este grupo de sujetos que puntuó por encima de 84 en la prueba ECO 1. Los resultados muestran que tanto para el GE ($n = 19$) como para el GC ($n = 23$) no se encuentran diferencias estadísticamente significativas entre sus puntuaciones iniciales y finales. La ausencia de diferencias significativas se da para ambos grupos en la puntuación total y también en el resto de las dimensiones contempladas en la prueba. Para el GE, la puntuación total en el pretest fue de 89,514 (DT = 6,176), y la puntuación posttest, 86,917 (DT = 10,519); el valor t para muestras relacionadas arrojó un resultado de 1,442, con una significación asociada de 0,166 para 18 grados de libertad. Los resultados de las distintas dimensiones muestran puntuaciones iniciales en torno a los noventa puntos y similares para el posttest, siempre con ligeros decrementos. Por su parte, el GC obtuvo una puntuación pretest de 91,403 (DT = 6,010) y una posttest de 92,963 (DT = 12,440); el valor t fue de -0,552 y la significación de 0,586 para 22 grados de libertad. En cuanto a las dimensiones de la prueba, el comportamiento fue similar al descrito para el GE, si bien con puntuaciones un poco por encima, es decir, puntuaciones iniciales en torno a los noventa y tres puntos y finales que no guardan diferencia digna de mención con las primeras. En definitiva, tanto en uno como en otro grupo, aquellos alumnos con dificultades que puntuaron inicialmente por encima de los 84 puntos en la prueba de comprensión (en la mayoría de las ocasiones bastante por encima; el promedio para ambos grupos se situaba en torno a los noventa puntos en la prueba inicial), no lograron incrementar sus puntuaciones en la fase final del estudio. Las razones parecen tener que ver con el análisis que efectuábamos anteriormente relativo al grado de dificultad de la prueba en relación al margen posible de mejora de este grupo de alumnos, así como a la posible falta de

eficacia mostrada por el dispositivo EDPL en las condiciones experimentales en que fue aplicado, para el grupo de alumnos citado.

(c) *Criterio del profesorado + puntuación inicial en la prueba ECO 1+ puntuación inicial en Cattell*

Para la realización de los contrastes que respondan a este tercer criterio hemos seleccionado de entre los datos aquellos sujetos que fueron catalogados por sus profesores u orientadores con dificultades de aprendizaje en la lectura, al tiempo que obtenían puntuaciones que se situaban por debajo de una desviación típica de la media muestral, tanto para la prueba de comprensión ECO 1, como para la prueba de factor «G» de Cattell. Los resultados para el GE y el GC se muestran en la siguiente tabla.

Tabla 36. Comparación pretest-postest para la prueba ECO 1 en función del criterio de designación del profesorado más la puntuación inicial en la prueba de comprensión (puntuación ≤ 84) más la puntuación inicial en el test de inteligencia (puntuación ≤ 84)

Prueba t para muestras relacionadas

Gr.	Dimensiones	N	Pretest	DT	Postest	DT	T	Gl	Sign.
GE GC	Puntuación total	7 6	77,305 76,450	4,795 4,176	84,033 83,553	2,811 13,513	-2,646 -1,343	6 5	0,038* 0,237
GE GC	Concepción de la lectura	7 6	87,957 84,566	8,973 0,000	91,349 88,522	11,584 9,691	-1,000 -1,000	6 5	0,356 0,363
GE GC	Metaconocimiento estratégico	7 6	76,619 80,602	3,576 5,037	87,135 87,294	7,153 16,068	-3,267 -0,889	6 5	0,017* 0,415
GE GC	Integración textual	7 6	85,616 81,079	12,490 10,013	85,616 88,408	6,971 11,033	0,000 -1,168	6 5	1,000 0,296
GE GC	Integración texto-conocimientos	7 6	80,118 77,204	8,889 2,629	85,639 84,719	6,885 13,510	-1,279 -1,472	6 5	0,248 0,201
GE GC	Supervisión y autoevaluación	7 6	85,732 88,926	7,170 10,953	91,481 86,690	9,257 7,347	-1,441 0,415	6 5	0,200 0,695

El análisis estadístico de los datos nos ha revelado con relación al GE, que también para el alumnado que puntuaba bajo en la prueba de CI se mantienen las diferencias significativas entre las puntuaciones globales de comprensión en las fases pretest y postratamiento del estudio. Estas diferencias son importantes para la práctica totalidad de las dimensiones consideradas y llegan a ser también significativas al nivel de 0,05 para la dimensión de *Metaconocimiento estratégico*. En cambio, para el GC, a pesar de que la diferencia entre las puntuaciones medias es similar a la del GE, el estadístico no ofrece significatividad.

Hemos realizado un contraste entre las ganancias obtenidas por ambos grupos para evaluar las posibles diferencias de magnitud. Los resultados para la puntuación total y para las distintas dimensiones de la prueba no arrojan ningún valor de diferencia estadísticamente significativa. La dimensión en la que se observa una mayor distancia entre las puntuaciones medias, *Supervisión y autoevaluación*, obtiene un valor t de -1,214 con una significación asociada de 0,250 para 11 grados de libertad.

Al igual que hicimos cuando comparamos las puntuaciones pre y postest para los sujetos con bajas puntuaciones en ECO 1, nos interesa saber quienes han conformado en esta ocasión los grupos experimental y control. Evidentemente, dado que cumplen los tres criterios señalados al comienzo del apartado, conocemos que se trata de alumnos y alumnas que son considerados por el sistema educativo (profesorado y/o orientadores de los centros) como sujetos con dificultades de aprendizaje de la lectura, que, así mismo, han confirmado con sus bajas puntuaciones en una prueba de comprensión de textos esta valoración, y que obtienen igualmente puntuaciones bajas en una prueba de CI; aún así, queremos saber qué circunstancias y condiciones de apoyo acompañan al grupo de sujetos que, en el caso del GE, ha visto incrementadas sus puntuaciones de forma significativa entre las fases inicial y final del estudio. Por una parte, En lo que respecta al GE, la inspección de la matriz de datos nos muestra que este grupo está formado por 3 alumnos con discapacidad, de los cuales 2 presentan desventaja socio-cultural, recibiendo los tres AI; otros 2 alumnos con dificultades de aprendizaje y privación social participan de AF; y, por último, 2 alumnos con dificultades de aprendizaje reciben AI. Como ya ocurría cuando realizamos la comparación para el segundo de los criterios señalados, el alumnado que ha integrado este grupo se caracteriza especialmente por presentar dificultades de aprendizaje de la lectura que son consideradas severas, ya que ocasionan la asistencia al aula de apoyo a la integración, y acompañadas de desventaja social o discapacidad en la mayoría de los casos. Por otra parte, el GC está formado también por alumnos y alumnas con dificultades de aprendizaje acompañadas de otros problemas que inciden directa o indirectamente en el rendimiento escolar; de esta forma, 3 alumnos presentan discapacidad, uno de los cuales también presenta desventaja social, recibiendo los tres AI; 2 alumnos más presentan privación social, recibiendo uno de ellos AI y participando el otro de AF; por último, 1 alumno fue catalogado solo DA, recibiendo igualmente AI.

Nos interesaba también averiguar cómo se comportaron los alumnos que obtuvieron bajas puntuaciones en ECO 1, pero que, sin embargo, puntuaron por encima de 84 en la prueba de CI. Recordemos que habíamos considerado en nuestro proyecto una categoría de alumnos con DA que debía responder a un perfil caracterizado esencialmente por la ausencia de otros problemas relativos al aprendizaje que pudieran explicar la existencia de las dificultades lectoras, obteniendo puntuaciones significativamente más bajas que sus iguales en pruebas de lectura y puntuando en cambio dentro de los parámetros normales en las pruebas de CI. Este grupo de alumnos, decíamos, manifestaba especialmente sus problemas de lectura en los procesos relacionados con la decodificación o el reconocimiento de palabras. Hemos realizado los cálculos pertinentes para buscar —de forma independiente a la valoración que realizó el profesorado sobre este mismo asunto— a estos alumnos entre la muestra estudiada. Las puntuaciones observadas para este reducido grupo de alumnos ($n = 4$) en el GE reflejan ganancias entre las fases inicial y final del estudio, aunque sin llegar a ser significativas. La media pretest para la puntuación total fue de 78,506 (DT = 4,973), mientras que la puntuación postest fue de 83,553 (DT = 7,438); el valor t obtenido para muestras relacionadas fue de -0,827 y la significación asociada de 0,469 para 3 grados de libertad. En el resto de dimensiones las ganancias se mantuvieron constantes, especialmente para *Concepción sobre la lectura* (pretest: 90,501 / postest: 96,436), *Integración textual* (pretest: 81,690 / postest: 89,019) e *Integración texto-conocimientos previos* (pretest: 81,498 / postest: 87,939). En cambio, para la dimensión de *Supervisión y Autoevaluación*, los resultados reflejan pérdidas, sin ser significativas, en vez de ganancias. Como apuntábamos un poco más arriba, se había pedido al profesorado que

valorara qué tipo de dificultades de aprendizaje de la lectura presentaban los alumnos; en este sentido, hemos contrastado los resultados obtenidos con la valoración efectuada por los profesores para estos cuatro sujetos. La coincidencia es del 50%, es decir: para dos de estos sujetos, que habían sido catalogados por sus profesores como alumnos que presentaban dificultades en la decodificación y en la comprensión global de los textos conjuntamente (tipología denominada *variedad jardín*; Sánchez, 1999), las puntuaciones en ECO 1 y en Cattell han confirmado estas valoraciones; en cambio, para otros dos sujetos, la valoración de sus profesores como alumnos con especiales dificultades en la comprensión global, no se ha visto reflejada por las puntuaciones obtenidas por estos dos alumnos.

En el GC solo dos alumnos poseían las características que buscábamos en este contraste. Estos dos alumnos obtuvieron igualmente ganancias entre las fases inicial y final del estudio, tanto en la puntuación total como en el resto de las dimensiones contempladas. La información que teníamos sobre la valoración que había efectuado uno de los profesores colaboradores acerca del tipo de dificultades de aprendizaje relacionadas con la lectura para uno de estos alumnos, nos informaba de la ausencia de coincidencia entre esta valoración, que apuntaba a dificultades de comprensión global de los textos, y los resultados obtenidos de sus puntuaciones.

Las DDP en la prueba de comprensión ECO 1

Hemos adoptado un instrumento exploratorio creado por Mora (1998) con el objetivo de paliar algunas limitaciones de los recursos estadísticos habitualmente empleados en la investigación. Estas limitaciones podrían concretarse en: (a) tendencias de subgrupos dentro del grupo completo de sujetos, y (b) casos singulares que se apartan de la tendencia global del grupo. Este instrumento dirige su atención hacia la *dirección de las diferencias ponderadas* (DDP), otorgando la posibilidad de una valoración matizada de los resultados. Compartiendo condiciones con los diseños *intrasujeto*, cada elemento de la muestra actúa como su propio control. En este sentido, de igual forma que en los estudios citados de Mora, se estableció qué sujetos de cada grupo habían obtenido puntuaciones a favor de la hipótesis experimental, es decir puntuaciones que reflejaran una ganancia postratamiento, y qué sujetos habían obtenido puntuaciones en contra de la hipótesis, o lo que es lo mismo, ausencia de diferencias pretest-postest o pérdidas. Los resultados podrían matizar de este modo la comparación de medias simples pretest y postest, ya que en los contrastes estadísticos habitualmente utilizados, las ganancias observadas para un grupo son atribuidas al grupo en su conjunto y, a efectos del estudio, afectan al grupo completo; en cambio, podrían no reflejar el comportamiento de todos sus integrantes. Hemos tratado de profundizar anteriormente en el perfil de los sujetos que han obtenido ganancias en función de algunos de los criterios formulados. El instrumento exploratorio elaborado por Mora trata de corregir un posible problema relativo a la magnitud de ganancias o pérdidas registradas en cada grupo; para ello, como apuntábamos en líneas anteriores, en primera instancia se calcula el número de sujetos que puntúan *a favor de la hipótesis* experimental, considerando que esta condición se cumple cuando las puntuaciones postest superan a las pretest; en caso contrario, es decir, cuando las puntuaciones pretest superan a las finales, se considera que los sujetos puntúan *en contra de la hipótesis*. Para los casos en que la diferencia de medias es cero, se considera que, al no cumplir la condición de mejora postratamiento, corresponden a las puntuaciones en contra de hipótesis. Posteriormente, el número de puntuaciones a favor y en contra de hipótesis son expresadas mediante porcentajes,

mostrando de una forma más ilustrativa los resultados obtenidos, tomando en consideración el comportamiento de los integrantes de cada grupo contrastado. Sin embargo, a pesar de que los porcentajes simples obtenidos con el instrumento consiguen matizar y enriquecer los resultados, no incluyen aún la magnitud de las diferencias pretest que los sujetos consiguieron, por lo que posteriormente el instrumento añade a la información ofrecida por los porcentajes de sujetos ganadores o perdedores, la comparación entre sujetos una vez que se ha efectuado la suma de sus puntuaciones de ganancias o pérdidas a nivel individual. Esta comparación con las DDP enriquece la información anterior, obteniendo una nueva puntuación porcentual que incluye la magnitud de las diferencias obtenidas en las puntuaciones inicial y final para cada sujeto. Por último, se aplica también la *t* de Student para establecer la probabilidad asociada a la puntuación obtenida.

A pesar de que para los contrastes en función del primero de los criterios formulados, no obtuvimos diferencias significativas favorables a nuestra hipótesis, hemos realizado los cálculos de las DDP para este primer grupo, aunque eliminando del mismo aquellos sujetos que no precisaban atención educativa especial (Grupo reducido). Recordemos que para el GE el número de sujetos era de 28, mientras que eran 15 los que cumplían estas condiciones en el GC1. Los resultados del GE muestran que 12 sujetos obtuvieron pérdidas y 3 una diferencia nula entre el pretest y el postest, mientras que para 13 alumnos la puntuación final fue superior a la inicial. Los resultados porcentuales se recogen a continuación.

	Ganancias	Pérdidas	A favor de hipótesis %	En contra de hipótesis %
Sujetos	13	15	46,43	53,57
Puntos	87,47	90,47	49,16	50,84

Podemos observar que prácticamente no hay diferencias entre las puntuaciones de ganancias y de pérdidas; de hecho, los tres puntos de diferencia a favor de la columna de pérdidas son debidos a los puntos otorgados a las comparaciones que obtuvieron una diferencia nula entre pretest y postest. De no ser así, los porcentajes serían idénticos a favor y en contra de hipótesis. Lógicamente, el hecho de que esto ocurra es en sí mismo negativo con relación a la hipótesis experimental, e indica la ausencia de un efecto significativo del tratamiento aplicado.

En cuanto al GC1 recordemos que las diferencias pretest-postest se acercaban a la significatividad; veamos cómo ha sido la evolución de sus ganancias y pérdidas mediante los cálculos con las DDP.

	Ganancias	Pérdidas	A favor de hipótesis %	En contra de hipótesis %
Sujetos	8	7	53,33	46,67
Puntos	134,571	30,157	81,69	18,31

A pesar de que, como podemos ver, la diferencia de porcentajes de sujetos cuyas puntuaciones evolucionaron a favor y en contra de hipótesis no es grande, sí lo es la diferencia en sus puntuaciones porcentuales una vez que se ha incluido en el contraste la magnitud de sus ganancias o pérdidas. De esta forma podemos concluir que las

diferencias obtenidas para las ganancias han superado con creces a las diferencias mostradas cuando se han obtenido pérdidas

Nos queda ahora establecer la probabilidad con la que pueden ser obtenidos estos resultados. Tal y como apuntábamos en líneas anteriores, la prueba t de Student puede otorgarnos la posibilidad de contrastar las diferencias promedio obtenidas por los sujetos que puntuaron a favor o en contra de la hipótesis. Mora (1998), citando a Dixon y Massey (1969), recoge entre los requisitos para la aplicación de esta prueba, la distribución normal de la variable en la población y la medición en escala de intervalos. En nuestro caso, ya comprobamos que las puntuaciones obtenidas para la aplicación piloto de la prueba ECO 1, donde estaban también incluidos sujetos con dificultades de aprendizaje de la lectura, presentaban una distribución normal, por lo que a tenor de estos resultados la variable *comprensión de textos* estaría distribuida normalmente entre la población; no obstante hemos vuelto a realizar el contraste de normalidad mediante la prueba Kolmogorov-Smirnov para el grupo total de sujetos con dificultades de aprendizaje que realizaron las pruebas pretest. En este caso el valor del estadístico es de 0,101 y la significación asociada de 0,083 para 68 grados de libertad. Aunque este resultado no muestra significatividad para la distribución normal de las puntuaciones en este grupo de alumnos con dificultades, lo cual era, por otra parte, esperable en función del sesgo muestral que supone el hecho de que todos los sujetos de esta submuestra fueran catalogados previamente con problemas de aprendizaje en la lectura, hemos optado por la aplicación de la prueba t en virtud del análisis anterior realizado para la aplicación piloto, considerando que la muestra es más representativa y que, por consiguiente, valora de forma más óptima la distribución poblacional de la variable *comprensión de textos*. Por otra parte, la prueba t de Student presenta una serie de ventajas que resultan de especial interés en nuestro caso. En primer lugar puede ser aplicada con grupos no muy amplios; no requiere la igualdad de varianzas entre las puntuaciones pre y postest; y, lo que es más importante, no requiere que las puntuaciones pretest y postest sean independientes. Los resultados de los contrastes realizados para los grupos experimental y control se muestran en las tablas siguientes.

*Tabla 37. Contraste en función de las puntuaciones DDP para el GE en la prueba ECO 1
Prueba t para muestras independientes*

	N	Diferencia media (pre — post)	DT	Prueba t	gl.	Sign.
Grupo experimental	13 (G) 15 (P)	6,728 6,031	6,852 5,055	-0,309	26	0,760

En el contraste con la prueba t podemos ver que la diferencia de medias promedio en el GE es superior para las ganancias, pero esta diferencia es exigua y no refleja ninguna significatividad estadística.

*Tabla 38. Contraste en función de las puntuaciones DDP para el GC1 en la prueba ECO 1
Prueba t para muestras independientes*

	N	Diferencia media (pre — post)	DT	Prueba t	gl.	Sign.
Grupo control	8 (G) 7 (P)	16,821 4,308	12,516 3,589	-2,703	8,293	0,026*

El contraste con la prueba t para muestras independientes, para el que no se asumió la igualdad entre las varianzas, confirma las diferencias pretest-postest obtenidas por el GC1, a las que hacíamos mención anteriormente, resultando significativo y matizando la información obtenida con las medias simples de las dos mediciones.

Resultados en la prueba de comprensión de textos ECOS para educación secundaria

Abordamos igualmente para la muestra de secundaria nuestros análisis psicométricos en función de los distintos criterios que apuntamos al comienzo del capítulo. Igualmente, nos detendremos en analizar los resultados para los distintos subgrupos considerados, tanto en función del grado de dificultad presentado por los sujetos, como en función de los aplicadores. Comenzaremos, al igual que hicimos con la prueba para primaria, por analizar los resultados para el grupo completo de alumnos, en función de la valoración que realizó el profesorado colaborador acerca de la condición de alumno con dificultades de aprendizaje relacionadas con la lectura.

(a) Criterio de designación previa del profesorado sobre las dificultades de aprendizaje

Tabla 39. Comparaciones pretest-postest para la prueba ECOS en función del criterio de designación del profesorado de las dificultades de aprendizaje

Prueba t de muestras relacionadas

Gr.	Dimensiones	N	Pretest	DT	Postest	DT	t	gl	Sign.
GE GC	Puntuación total	25 34	87,446 90,922	9,733 7,670	95,375 96,071	12,136 11,525	-5,338 -3,887	24 33	0,000** 0,000**
GE GC	Concepción de la lectura	25 34	90,856 93,448	9,261 12,648	98,777 98,544	14,503 15,215	-2,551 -1,559	24 33	0,018* 0,128
GE GC	Metaconocimiento estratégico	25 34	90,693 92,280	10,802 13,606	97,072 97,948	10,768 13,697	-3,588 -2,650	24 33	0,001** 0,012*
GE GC	Conocimientos previos	25 34	86,775 92,430	12,394 11,457	94,816 97,536	9,407 12,728	-3,029 -2,450	24 33	0,006** 0,020*
GE GC	Integración textual	25 34	91,098 92,309	12,708 8,552	95,333 95,706	13,574 12,856	-1,660 -1,659	24 33	0,110 0,107
GE GC	Integración texto-conocimientos	25 34	88,782 91,201	9,624 8,214	95,440 95,699	14,111 12,772	-3,252 -2,717	24 33	0,003** 0,010**
GE GC	Supervisión y autoevaluación	25 34	88,032 95,007	6,723 10,044	98,324 97,530	15,324 11,737	-3,674 -0,984	24 33	0,001** 0,332

Una primera inspección de los datos recogidos en la tabla nos muestra las diferencias muy significativas para el GE entre las fases pretest y postratamiento del estudio. Estas diferencias son significativas al nivel 0,01 para la puntuación total de la prueba, así como para las dimensiones *Metaconocimiento estratégico*, *Conocimientos previos*, *Integración texto-conocimientos previos* y *Supervisión y autoevaluación*; presentando una significación al nivel 0,05 para la dimensión *Concepción de la lectura*, y no resultando significativa a pesar de las diferencias observadas para la dimensión *Integración textual*. Estos resultados quedan velados en parte por el hecho de que, como podemos observar, también el GC ha obtenido resultados que muestran diferencias muy significativas entre las fases inicial y final. Las diferencias observadas para el GC, además de en la puntuación total, resultan significativas al 0,01 en la dimensión

Integración texto-conocimientos previos y al nivel de 0,05 en *Metaconocimiento estratégico* y en *Conocimientos previos*; en cambio no resultan significativas para el resto de dimensiones contempladas. Podemos igualmente constatar que las diferencias obtenidas por el GE resultan ser mayores que las obtenidas por el GC; hemos realizado un contraste entre las diferencias promedio de las medias de los grupos experimental y control para comprobar la posible significatividad de estas diferencias observadas. Los resultados para la puntuación total muestran un valor t para muestras independientes de -1,388, con una significación asociada de 0,170 para 57 gl. Por tanto, a pesar de que en el GE se observan mayores diferencias, éstas no llegan a ser significativas con respecto al GC1. En cuanto a las distintas dimensiones contempladas, sí hemos hallado diferencias significativas a favor del GE entre las ganancias medias obtenidas por ambos grupos para *Supervisión y autoevaluación del proceso de comprensión*, con un valor t de -2,026 y una significación de 0,047 para 57 gl.

Al igual que hicimos con la prueba de primaria, hemos vuelto a realizar las comparaciones excluyendo de la muestra a los sujetos que no recibían AI o no presentaban discapacidad o deprivación socio-cultural; en este caso, hemos añadido la condición de no participar de programas de DC para ser también excluido de esta comparación. Las razones de este nuevo contraste estriban en que también para la muestra de secundaria podían observarse diferencias en la conformación de los grupos en función del grado aparente de sus problemas de aprendizaje, así como de la atención educativa que recibían, lo que indirectamente puede llevarnos al grado de dificultad. Esta presencia diferencial de sujetos con menores dificultades se hacía patente en la comparación entre alumnos de uno y otro grupo que no recibían AI ni planteaban problemas de discapacidad o deprivación social, siendo este número de 18 en el GC1 y de tan solo 8 en el GE, habiendo incluido a los 8 sujetos que participan de programas de DC en este subgrupo. También es diferente el número de alumnos que solo fueron catalogados con DA por el profesorado, en ausencia de otros problemas de aprendizaje, siendo de 24 en el GC1 y de 17 en el GE. Por el contrario, los alumnos con discapacidad eran más en el GC1 (10) que en el GE (6). En definitiva, hemos querido saber qué resultados se darían si comparáramos sólo los sujetos que respondían a una concepción más restrictiva de las dificultades de aprendizaje. Estos resultados muestran que el GE no varía sus efectivos, por lo que el contraste arroja idénticos resultados que los expresados en la tabla. Sin embargo, el GC1, como ya ocurriera en primaria, ve reducido ampliamente el número de sujetos que intervienen en la nueva comparación, pasando de 34 a 16. Los resultados, sin embargo, siguen mostrando diferencias muy significativas para la puntuación total en este grupo reducido, obteniendo una media inicial de 88,199 (DT = 6,448) y una final de 93,353 (DT = 8,217), con un valor t de -4,392 y una significación de 0,001 para 15 gl. En cuanto al resto de dimensiones, las diferencias siguen siendo significativas al nivel 0,05 para dos de ellas: *Metaconocimiento estratégico* y *Conocimientos previos*.

Cuando analizábamos la equivalencia en función del sexo entre los grupos experimental y control, mostrábamos que para el grupo de educación secundaria había una mayor presencia de niños en el GC que resultaba estadísticamente significativa, pudiendo constituir esta circunstancia una desventaja en el caso de que los niños obtuvieran finalmente puntuaciones significativamente más bajas que sus compañeras. Hemos realizado un contraste para verificar esta posibilidad. Los resultados se muestran en la siguiente tabla.

Tabla 40. Comparaciones pretest-postest por sexos en la prueba ECOS (puntuación total)
Prueba t para muestras relacionadas

Grupos		N	Pretest	DT	Postest	DT	valor t	gl	Significación
GE	niñas	9	91,896	11,599	98,808	14,082	-2,304	8	0,050*
	niños	16	84,944	7,826	93,443	10,901	-5,122	15	0,000**
GC1	niñas	12	89,043	6,058	94,227	9,672	-2,675	11	0,022*
	niños	22	91,947	8,372	97,076	12,519	-2,873	21	0,009**

Los resultados muestran la existencia de diferencias significativas en todos los contrastes realizados; las diferencias más acusadas se dan precisamente entre las puntuaciones pretest y postest de los niños, aunque esto puede deberse al mayor número de éstos en los grupos de comparación. En cualquier caso, podemos afirmar que la mayor presencia de niños en el GC1 no parece haber ocasionado una desventaja; antes bien, podemos ver que sus resultados reflejan una mayor diferencia entre las puntuaciones iniciales y finales que la que obtienen sus compañeras. Por otra parte, hemos realizado previamente un contraste entre las puntuaciones iniciales de niños y niñas, tanto para GE como para GC con el objetivo de comprobar si existían diferencias significativas de partida. Los resultados mostraron ausencia de diferencias significativas para el GC y también para el GE (valor $t = 1,604$; sig. 0,134 para 23 grados de libertad), a pesar de que las niñas partían con una puntuación media inicial que superaba en más de seis puntos a la de los niños.

Contrastes en función de los distintos subgrupos considerados en la muestra

Por otra parte, nos interesa conocer los resultados obtenidos por los distintos subgrupos que hemos considerado en nuestro estudio. Recordemos que, a diferencia de lo que ocurría en primaria, la presencia de alumnos y alumnas que recibían apoyo a la integración estaba más equilibrada en esta etapa. En la siguiente tabla recogemos los resultados obtenidos para los grupos experimental y control tras los análisis efectuados para estos subgrupos. Aunque hemos reflejado tan sólo las puntuaciones totales obtenidas por los grupos, debemos destacar algunos datos con relación a las dimensiones contempladas. Así, para el contraste en función de las dificultades de aprendizaje en ausencia de discapacidad o privación social, el GE obtuvo diferencias muy significativas (al nivel 0,01 de significación) para las dimensiones *Conocimientos previos*, *Integración texto-conocimientos* y *Supervisión y autoevaluación del proceso de comprensión*, además de para la puntuación total; y diferencias significativas al nivel de 0,05 para la dimensión de *Metaconocimiento estratégico*. Por su parte, el GC1, como podemos observar, también mostró diferencias muy significativas entre las puntuaciones totales pretest y postest, aunque en los contrastes posteriores sólo obtuvo diferencias significativas, al 0,05, en la dimensión *Integración texto-conocimientos*. Para el contraste en función de la discapacidad, el GE obtuvo también diferencias favorables en todas las dimensiones contempladas, llegando éstas a ser significativas (nivel 0,05) para *Metaconocimiento estratégico*. De igual forma, el GC1 obtuvo diferencias favorables al postest en la práctica totalidad de las dimensiones, llegando a

Tabla 41. Comparaciones pretest-postest en la prueba ECOS (puntuación total) para los distintos subgrupos contemplados en la muestra (DA = dificultades de aprendizaje; DIS = discapacidad; AI = recibe apoyo a la integración; DEP = privación socio-cultural; RL = recibe refuerzo de Lengua; NA. = no recibe apoyo; DC = diversificación curricular)

Prueba t para muestras relacionadas

Grupos	N	Pretest	DT	Postest	DT	valor t	gl	significación
GE DA	17	89,986	9,042	99,432	11,714	-4,796	16	0,000**
GC1 DA	24	91,876	7,900	97,120	12,340	-3,023	23	0,006**
GE DIS	6	79,518	6,339	85,064	7,343	-3,005	5	0,030*
GC1 DIS	10	88,633	6,923	93,552	9,377	-2,666	9	0,026*
GE AI	17	82,922	5,853	91,177	10,451	-4,429	16	0,000**
GC1 AI	16	88,199	6,448	93,353	8,217	-4,392	15	0,001**
GE DEP	2	89,645	16,367	91,816	11,252	-0,600	1	0,656
GC1 DEP	0	---	---	---	---	---	---	---
GE NA, RL, DC	8	97,060	9,547	104,294	11,029	-2,804	7	0,026*
GC NA, RL, DC	18	93,343	8,020	98,487	13,606	-2,224	17	0,040*

la significatividad, también con un nivel inferior al 0,05, para la dimensión *Conocimientos previos*. En cuanto a la variable AI, el GE mostraba diferencias muy favorables al postest en todas las dimensiones; entre ellas llegaban a ser significativas al nivel 0,01 para *Conocimientos previos*, y al nivel 0,05 para *Metaconocimiento estratégico* e *Integración texto-conocimientos*. El GC1 se comportó de forma similar, obteniendo diferencias significativas (nivel 0,05) para las dimensiones *Metaconocimiento estratégico* y *Conocimientos previos*. Por otra parte, como podemos observar, el contraste para el alumnado que presentaba privación socio-cultural no ha podido llevarse a cabo para el GC1 debido a que, aunque había en la muestra dos alumnos bajo esta condición, éstos han sido incluidos en los contrastes para el subgrupo discapacidad al encontrarse igualmente valorados en este grupo —otros alumnos del GE se encontraban en circunstancias similares—. Por último, hacemos mención del alumnado valorado con dificultades de aprendizaje por el profesorado o por los orientadores colaboradores, pero que no recibía AI, y que, en secundaria suele participar, bien de la materia Refuerzo de Lengua (RL) en los cursos iniciales de la etapa, o bien de programas de Diversificación Curricular (DC) en el último curso. En nuestros grupos, 13 alumnos que recibían RL formaban parte del GE y 18 del GC; algunos de estos alumnos, además de recibir este tipo de refuerzo, poseían otras características (privación social, discapacidad o recibir también AI) que han ocasionado su inclusión en otros subgrupos para los contrastes realizados. Solo aquellos que no recibían AI y no estaban incluidos en las categorías de discapacidad o privación, han integrado por tanto esta última comparación. Por su parte, el alumnado que participaba de programas de DC se ha concentrado en el GE, con 8 alumnos, mientras que no había ninguno en el GC1. Al igual que comentábamos para los alumnos de RL, algunos de los que participaban de DC presentaban otras características relativas a sus dificultades que aconsejaron su inclusión en otros subgrupos. El contraste, aparte de mostrar diferencias significativas para la puntuación total de ambos grupos, muestra para el GE diferencias muy significativas (nivel 0,01) para la dimensión *Supervisión* y

autoevaluación, mientras que estas diferencias son significativas al 0,05 para *Integración texto-conocimientos*. Por su parte, el GC también muestra diferencias estadísticamente significativas al nivel 0,05 para la dimensión *Integración texto-conocimientos*.

A pesar de que podemos observar diferencias significativas para la puntuación total de ambos grupos en los diferentes contrastes realizados, éstas parecen ser de mayor intensidad en el GE. En lo referente a la posible significatividad de las diferencias de medias observadas entre las puntuaciones de los grupos experimental y control, hemos realizado un contraste para verificar esta posibilidad. De confirmarse la existencia de diferencias reseñables entre las puntuaciones obtenidas del cálculo de las diferencias de medias, podríamos confirmar mejores resultados para el GE a pesar de que el GC1 hubiera obtenido diferencias significativas entre sus puntuaciones pre y postest. Los resultados de estos contrastes para los diferentes subgrupos se muestran en la siguiente tabla. Las puntuaciones negativas, obtenidas de la diferencia entre la puntuación pretest y la postest, expresan ganancias, mientras que las positivas expresan pérdidas. Todos los contrastes han sido efectuados tras comprobar la existencia de homogeneidad de las varianzas.

Tabla 42. Contraste de diferencias de medias pretest-postest entre los grupos experimental y control en la prueba ECOS (puntuación total), en función de los distintos subgrupos contemplados en la muestra (DA = dificultades de aprendizaje; DIS = discapacidad; AI = recibe apoyo a la integración; DEP = privación socio-cultural; RL = recibe refuerzo de Lengua; NA. = no recibe apoyo)

Prueba t para muestras independientes

Grupos	N	Diferencia de medias	DT	valor t	gl	significación
GE DA GC1 DA	17 24	-9,446 -5,244	8,121 8,498	-1,588	39	0,120
GE DIS GC1 DIS	6 10	-5,545 -4,918	4,521 5,835	-0,225	14	0,825
GE AI GC1 AI	17 16	-8,254 -5,154	7,685 4,694	-1,388	31	0,175
GE DEP GC1 DEP	2 0	<i>No se han podido realizar estos cálculos</i>				
GE NA, RL GC1 NA, RL	8 18	-7,233 -5,143	7,295 9,815	-0,537	24	0,596

La inspección de los datos nos muestra que a pesar de que las diferencias promedio de las puntuaciones inicial y final son mayores en el GE, éstas no llegan a ser significativas estadísticamente utilizando los parámetros de significación habituales, lo que ya ocurriera cuando realizamos estos mismos cálculos para el grupo completo.

Queremos ahora fijar nuestra atención en las posibles diferencias que hayan podido existir en función del nivel educativo de los efectivos que han integrado los grupos. Habíamos realizado los análisis de equivalencia tomando en consideración como unidad de nivel educativo el ciclo, por lo que hemos calculado los estadísticos de contraste entre los grupos para los dos ciclos contemplados. Los resultados muestran diferencias

significativas para todos los contrastes realizados con las puntuaciones medias totales de la prueba. En el primer ciclo y para el GE, los doce alumnos que han integrado este subgrupo obtuvieron un promedio de 84,702 en el pretest (DT = 5,251) y de 93,383 (DT = 11,248) en el postest; el valor t observado fue de -3,332 y la significación asociada de 0,007 para 11 grados de libertad. Además, las diferencias fueron también muy significativas para la dimensión *Conocimientos previos* y significativas al nivel 0,05 para *Integración texto-conocimientos*. El GC1 por su parte (n = 25; recuérdese que se encontraron diferencias significativas para el número de integrantes de este ciclo a favor de este grupo) obtuvo un pretest de 90,687 (DT = 7,852) y un postest de 95,143 (DT = 11,581); el valor t fue de -2,691 y la significación de 0,013 para 24 gl. Este grupo también obtuvo diferencias significativas al nivel 0,05 en *Metaconocimiento estratégico*. En lo referente al segundo ciclo de secundaria, las puntuaciones del GE fueron 89,979 (DT = 12,235) en el pretest y de 97,213 (DT = 13,076) en el postest; t = -4,438; sig. 0,001 para 12 gl. Igualmente, este grupo obtuvo diferencias estadísticamente significativas al nivel 0,01 para las dimensiones *Metaconocimiento estratégico* y *Supervisión y autoevaluación*. Por último, para el GC1, las puntuaciones fueron de 91,574 (DT = 7,552) en la fase inicial y de 98,647 (DT = 11,636) en la fase postest, con un valor t de -3,599 y una significación asociada de 0,007 para 8 gl., no obteniendo este subgrupo diferencias que resultaran significativas para las dimensiones de la prueba.

Contrastes en función de los grupos de aplicación del dispositivo

Queremos ahora, al igual que hicimos con la prueba elaborada para primaria, abordar el análisis de los resultados para el GE en función de los aplicadores que han colaborado en la puesta en práctica experimental del dispositivo de evaluación dinámica. Ya hicimos mención cuando introducíamos el análisis de los resultados para primaria, que el objetivo principal de estos contrastes era conocer la posible incidencia que el factor personal del aplicador, concretado en aspectos como su estilo docente, su modo de interactuar con los sujetos, o la adecuación personal en el seguimiento de las pautas establecidas con carácter general, haya podido tener en los resultados obtenidos. En la siguiente tabla podemos observar las comparaciones por grupos a los que un mismo evaluador aplicó el dispositivo en el grupo experimental.

Observando los datos podemos comprobar que las diferencias significativas encontradas anteriormente en el contraste realizado para el grupo completo de alumnos con dificultades de aprendizaje de secundaria, se reflejan de un modo casi generalizado en los distintos grupos de aplicación del dispositivo. Tan solo el grupo J parece haber obtenido resultados más modestos, aunque obtuvo diferencias a favor de las puntuaciones postest y en una de las dimensiones éstas llegan a ser significativas. También el grupo H obtiene diferencias que no llegan a ser significativas en la puntuación total, aunque en este caso, las diferencias entre las puntuaciones pretest y postest son mayores. Por otra parte, es necesario tener en cuenta que el número de sujetos que componen cada uno de los subgrupos es muy pequeño, con lo que el rechazo de la hipótesis nula de igualdad entre las medias se hace muy difícil. Sin embargo, a pesar de lo exiguo de algunos de los grupos, en tres de ellos las diferencias han llegado a ser significativas o muy significativas en la puntuación total; además, en estos grupos (G, I y K), los resultados muestran diferencias también relevantes desde el punto de vista estadístico en algunas de las dimensiones contempladas, principalmente: *Conocimientos previos*, *Integración texto-conocimientos* y *Supervisión y autoevaluación*.

Tabla 43. Comparaciones pretest-postest para la prueba ECOS por grupos de aplicadores Estadísticos de muestras relacionadas

Grupos	Dimensiones		Media	N	DT	t	gl	Sign.
Grupo G IES Leonor de Guzmán	Puntuación total	Pretest	97,060	8	9,547 11,029	-2,804	7	0,026*
		Postest	104,294					
	Concepción sobre la lectura	Pretest	89,990	8	8,751 12,810	-1,528	7	0,170
		Postest	96,178					
	Metaconocimiento estratégico	Pretest	99,364	8	9,354 8,699	-2,198	7	0,064
		Postest	105,178					
	Conocimientos previos	Pretest	98,745	8	14,447 9,768	0,424	7	0,685
		Postest	96,461					
	Integración textual	Pretest	99,952	8	12,861 14,422	-0,801	7	0,450
		Postest	104,163					
	Integración texto-conocimientos	Pretest	97,982	8	10,601 11,262	-3,000	7	0,020*
		Postest	104,730					
	Supervisión y Autoevaluación	Pretest	89,962	8	8,104 13,740	-4,782	7	0,002**
		Postest	108,723					

Grupos	Dimensiones		Media	N	DT	t	gl	Sign.
Grupo H IES Leonor de Guzmán	Puntuación total	Pretest	79,880	4	5,332 7,957	-1,890	3	0,155
		Postest	83,497					
	Concepción sobre la lectura	Pretest	93,084	4	12,376 14,290	-1,000	3	0,391
		Postest	99,272					
	Metaconocimiento estratégico	Pretest	81,922	4	0,000 9,966	-1,667	3	0,194
		Postest	90,228					
	Conocimientos previos	Pretest	78,186	4	5,275 4,568	-3,000	3	0,058
		Postest	85,039					
	Integración textual	Pretest	83,109	4	9,215 11,371	-1,127	3	0,342
		Postest	89,125					
	Integración texto-conocimientos	Pretest	82,799	4	4,307 4,307	2,449	3	0,092
		Postest	78,300					
	Supervisión y Autoevaluación	Pretest	87,282	4	5,360 5,360	No se puede calcular t		
		Postest	87,282					

Grupos	Dimensiones		Media	N	DT	t	gl	Sign.
Grupo I IES Mariana Pineda	Puntuación total	Pretest	83,135	4	9,188 13,067	-6,195	3	0,008**
		Postest	96,517					
	Concepción sobre la lectura	Pretest	86,896	4	0,000 12,376	-1,000	3	0,391
		Postest	93,084					
	Metaconocimiento estratégico	Pretest	90,228	4	13,698 8,360	-0,577	3	0,604
		Postest	93,550					
	Conocimientos previos	Pretest	82,755	4	7,460 8,748	-7,000	3	0,006**
		Postest	98,745					
	Integración textual	Pretest	89,125	4	14,370 11,787	-3,000	3	0,058
		Postest	99,952					
	Integración texto-conocimientos	Pretest	81,674	4	5,807 14,225	-2,941	3	0,060
		Postest	97,419					
	Supervisión y Autoevaluación	Pretest	84,602	4	0,000 21,441	-1,000	3	0,391
		Postest	95,322					

Tabla 43. Comparaciones pretest-postest para la prueba ECOS por grupos de aplicadores (Continuación) Estadísticos de muestras relacionadas

Grupos	Dimensiones		Media	N	DT	t	gl	Sign.
Grupo J IES Torreblanca	Puntuación total	Pretest	84,727	5	3,000 5,732	-0,557	4	0,607
		Postest	86,463					
	Concepción sobre la lectura	Pretest	96,796	5	13,557 20,709	-0,784	4	0,477
		Postest	106,697					
	Metaconocimiento estratégico	Pretest	87,238	5	7,278 12,070	-1,177	4	0,305
		Postest	91,225					
	Conocimientos previos	Pretest	82,755	5	6,461 4,086	-4,000	4	0,016*
		Postest	90,065					
	Integración textual	Pretest	89,365	5	11,489 4,026	0,767	4	0,486
		Postest	84,553					
	Integración texto-conocimientos	Pretest	84,823	5	5,865 10,834	-0,389	4	0,717
		Postest	86,623					
	Supervisión y Autoevaluación	Pretest	88,890	5	9,588 5,871	0,000	4	1,000
		Postest	88,890					

Grupos	Dimensiones		Media	N	DT	t	gl	Sign.
Grupo K IES Torreblanca	Puntuación total	Pretest	83,497	4	6,292 7,779	-15,556	3	0,001**
		Postest	99,411					
	Concepción sobre la lectura	Pretest	86,896	4	0,000 14,290	-1,732	3	0,182
		Postest	99,272					
	Metaconocimiento estratégico	Pretest	86,906	4	11,347 8,578	-2,049	3	0,133
		Postest	98,534					
	Conocimientos previos	Pretest	80,470	4	4,568 8,748	-3,873	3	0,030*
		Postest	103,314					
	Integración textual	Pretest	85,515	4	9,624 14,437	-1,192	3	0,319
		Postest	92,734					
	Integración texto-conocimientos	Pretest	88,422	4	3,673 9,274	-3,434	3	0,041*
		Postest	103,043					
	Supervisión y Autoevaluación	Pretest	87,282	4	5,360 16,080	-2,324	3	0,103
		Postest	103,363					

El contraste general efectuado para comprobar la posible significatividad de las diferencias obtenidas por los diferentes grupos de aplicación, arroja un valor F de 4,740, con una significación asociada de 0,007 para 4 gl. intergrupos y 20 intragrupos. Por su parte, la prueba de homogeneidad de varianzas mostró que las varianzas de los grupos no diferían de forma significativa (Levene = 1,493; sig. 0,242). En los contrastes posteriores para comparaciones múltiples realizados por pares de grupos, la aplicación de los estadísticos de Bonferroni y Scheffé mostraron significatividad tan sólo entre las diferencias promedio de los grupos J y K (En la prueba de Bonferroni, diferencia de medias J-K = -14,178; sig. 0,017). En general, podríamos concluir sobre estas comparaciones, efectuadas en función de los aplicadores, que se han producido ganancias en todos los grupos de aplicación y que en tres de ellos estas ganancias son muy significativas. Así mismo, se producen diferencias significativas en cuanto a la magnitud de estas ganancias entre dos de los grupos de aplicación. Posteriormente, cuando abordemos la evaluación de la aplicación del dispositivo de evaluación dinámica EDPL desde una óptica cualitativa y complementaria a esta, podremos profundizar en las posibles relaciones establecidas entre los patrones metodológicos de evaluación dinámica puestos en práctica por los aplicadores, y los resultados obtenidos.

(b) Criterio del profesorado + puntuación inicial en la prueba ECOS

Nos disponemos ahora a realizar comparaciones entre los resultados obtenidos por los grupos experimental y control desde la adopción del segundo de los criterios expuesto al comienzo de este capítulo. Este criterio, como recordaremos, se basaba en la conformación de los grupos de estudio en función de la valoración del profesorado y de los orientadores, así como en función también de su puntuación inicial en la prueba de comprensión de textos. Hemos seleccionado los casos que cumplen esta condición y aplicado el estadístico de muestras relacionadas. Los resultados se muestran en la tabla siguiente.

Tabla 44. Comparación pretest-postest para la prueba ECOS en función del criterio de designación del profesorado más la puntuación inicial en la prueba (puntuación ≤ 84)

Prueba t para muestras relacionadas

Gr.	Dimensiones	N	Pretest	DT	Postest	DT	t	gl	Sign.
GE GC	Puntuación total	12 7	79,398 80,965	3,173 2,641	88,078 89,025	7,362 6,298	-5,196 -2,740	11 6	0,000** 0,034*
GE GC	Concepción de la lectura	12 7	88,958 86,896	7,145 0,000	93,084 86,896	11,194 0,000	-1,000	11	0,339
							No se puede calcular t		
GE GC	Metaconocimiento estratégico	12 7	83,030 80,024	4,769 3,242	91,889 91,415	8,732 11,417	-4,000 -2,521	11 6	0,002** 0,045*
GE GC	Conocimientos previos	12 7	79,709 82,755	5,951 5,275	92,654 90,587	9,901 16,199	-4,214 -1,279	11 6	0,001** 0,248
GE GC	Integración textual	12 7	82,307 84,140	6,269 6,641	87,520 91,703	9,281 7,717	-1,680 -1,577	11 6	0,121 0,166
GE GC	Integración texto-conocimientos	12 7	82,424 85,209	6,166 6,208	89,172 89,708	10,651 8,095	-3,200 -1,146	11 6	0,008** 0,296
GE GC	Supervisión y autoevaluación	12 7	84,602 87,665	0,000 5,231	89,962 92,259	9,697 8,104	-1,915 -1,162	11 6	0,082 0,289

El contraste realizado bajo este segundo criterio nos muestra que las ganancias obtenidas en el postest para la puntuación total con relación a la media inicial, resultan significativas estadísticamente, tanto para el GE como para el GC, aunque la significación asociada a las diferencias encontradas en el GE exprese un mayor nivel de confianza. Por otra parte, podemos observar que en el resto de dimensiones, las diferencias mostradas por el GE resultan muy significativas para *Metaconocimiento estratégico*, *Conocimientos previos* e *Integración texto-conocimientos*. El GC, por su parte, muestra ganancias significativas al nivel 0,05 en la dimensión *Metaconocimiento estratégico*. Estas diferencias entre las ganancias obtenidas por ambos grupos podrían ser relevantes desde el punto de vista estadístico, lo que fue analizado mediante un contraste entre las diferencias promedio pretest-postest obtenidas por uno y otro grupo. En este sentido, la constatación de que las ganancias fueran significativamente distintas nos ofrecería información relevante acerca de los procesos implicados en la lectura sobre los que habría actuado decisivamente el dispositivo aplicado. Los resultados de estos contrastes no nos ofrecieron en cambio diferencias estadísticamente significativas entre las ganancias obtenidas por los sujetos de ambos grupos en cada una de las dimensiones. Sin bien, en algunas comparaciones las diferentes magnitudes de las ganancias pueden ser importantes, no resulta suficiente como para reflejarse en los

contrastes estadísticos. La comparación que más cerca está de la significación estadística es la que se produce para la dimensión *Concepción sobre la lectura*, en la que, sin asumir varianzas iguales, el valor t para muestras independientes arroja un valor de -1,000, con una significación asociada de 0,339 para 11,00 grados de libertad.

Hemos querido saber quien ha integrado estos subgrupos de sujetos que puntuaron bajo en la prueba de comprensión, y si podían estimarse diferencias en cuanto a su tipología o condiciones de apoyo que pudieran complementar los resultados. Al igual que hicimos con la prueba de primaria, hemos rastreado en nuestra matriz de datos para comprobar qué alumnos estaban detrás de estos contrastes. Los resultados nos muestran que en el GE, 11 de los 12 sujetos que integraban este subgrupo recibían apoyo a la integración, de los cuales 5 presentaban discapacidad, 2 desventaja socio-cultural y 7 recibían también refuerzo de Lengua; por último, el alumno restante presentaba desventaja socio-cultural y participaba del programa de diversificación curricular. En cuanto al GC, 5 de los integrantes precisaban apoyo a la integración, de los cuales 2 presentaban discapacidad y otros dos recibían refuerzo de Lengua; los dos alumnos restantes habían sido catalogados por sus profesores sólo con dificultades de aprendizaje en la lectura y no recibían ningún tipo de apoyo educativo.

Hemos realizado nuevos cálculos sin tener ahora en cuenta a estos alumnos que, aunque han obtenido una puntuación baja en la prueba criterio, podrían presentar un menor grado de dificultad a tenor de los análisis que hemos revisado, y que apuntaban igualmente a los menores resultados de aquellos que precisaban de una atención educativa más específica. Los análisis efectuados nos muestran que las ganancias obtenidas, tanto por los sujetos que precisaban de condiciones de apoyo más específicas, como de aquellos que no requerían en principio de atención especializada, resultaron significativas, sin que se observaran mayores diferencias que la puntuación inicial de partida que sí resultó significativamente mayor para el grupo que no recibía AI ($t = 2,044$; sig.0,049 para 32 gl.).

Por lo que respecta a los sujetos que obtuvieron puntuaciones por encima de 84 en la prueba de comprensión, los integrantes del GE mostraron en sus resultados ganancias significativas para la puntuación total, obteniendo una puntuación media en el pretest de 94,876 (DT = 7,460) y de 102,110 (DT = 11,937) en el postest, con un valor t para muestras relacionadas de -2,944 y una significación asociada de 0,012 para 12 grados de libertad. Para el resto de dimensiones, las puntuaciones reflejaron diferencias significativas al nivel 0,01 para la dimensión *Supervisión y autoevaluación*, y para *Concepción sobre la lectura* al nivel 0,05. Por su parte, el GC obtuvo en la fase inicial una puntuación de 93,503 (DT = 6,280) y de 97,897 (DT = 11,942) en el postest, con un valor t de -2,976 y una significación asociada de 0,006 para 26 grados de libertad. En cuanto a las dimensiones, tan solo se mostraron significativas estadísticamente, al nivel 0,05, las diferencias para la dimensión *Integración texto-conocimientos*. Al igual que hicimos anteriormente para los sujetos que habían puntuado bajo en la prueba de comprensión, hemos realizado los contrastes de diferencias de ganancias medias entre estos dos subgrupos que obtuvieron puntuaciones más altas. Los resultados muestran esta vez diferencias significativas a favor del GE en la dimensión *Supervisión y autoevaluación del proceso de comprensión*. La diferencia observada entre las medias de ambos grupos ha sido amplia, de 12,858 puntos. El valor de la prueba t para muestras independientes, asumiendo la homogeneidad de las varianzas, ha sido de -2,374 y la significación asociada de 0,023 para 38 gl. Este resultado, que ya se produjo cuando

realizamos estos mismos cálculos en función del primero de los criterios contemplados, puede ser relevante a la hora de explicar cómo se ha producido el funcionamiento del dispositivo aplicado y en qué procesos de los trabajados ha podido ejercer una mayor influencia, máxime cuando se ha producido este resultado en una de las dimensiones directamente relacionadas con los procesos metacognitivos implicados en la comprensión.

(c) Criterio del profesorado + puntuación inicial en la prueba ECOS + puntuación inicial en Cattell

Hemos seleccionado para estas comparaciones a los sujetos que además de puntuar por debajo de una desviación típica de la media en la prueba de comprensión, obtuvieron también una puntuación similar en la prueba de CI. Los resultados nos otorgarán información precisamente sobre aquel grupo de alumnos que presenta mayores dificultades de entre la muestra del estudio. En la siguiente tabla podemos explorar las puntuaciones medias, obtenidas por los grupos experimental y control, así como el valor de los contrastes realizados.

Tabla 45. Comparación pretest-postest para la prueba ECOS en función del criterio de designación del profesorado más la puntuación inicial en la prueba de comprensión (puntuación ≤ 84) más la puntuación inicial en el test de inteligencia (puntuación ≤ 84)

Prueba t para muestras relacionadas

Gr.	Dimensiones	N	Pretest	DT	Postest	DT	T	Gl	Sign.
GE GC	Puntuación total	7 4	79,932 79,880	3,204 3,207	85,719 92,177	7,219 4,923	-3,098 -3,833	6 3	0,021* 0,031*
GE GC	Concepción de la lectura	7 4	90,432 86,896	9,355 0,000	93,968 86,896	12,077 0,000	-0,548	6	0,604
							No se puede calcular t		
GE GC	Metaconocimiento estratégico	7 4	83,821 80,261	3,242 3,322	90,465 93,550	8,329 11,347	-2,646 -1,852	6 3	0,038* 0,161
GE GC	Conocimientos previos	7 4	78,839 80,470	7,189 4,568	87,976 91,892	4,884 16,682	-4,583 -1,667	6 3	0,004** 0,194
GE GC	Integración textual	7 4	84,140 84,312	7,717 2,406	87,578 92,734	8,723 8,335	-0,853 -2,049	6 3	0,426 0,133
GE GC	Integración texto-conocimientos	7 4	81,996 83,924	5,101 3,673	86,494 95,170	11,573 5,807	-1,323 -3,873	6 3	0,234 0,030*
GE GC	Supervisión y autoevaluación	7 4	84,602 84,602	0,000 0,000	86,133 95,322	4,052 8,753	-1,000 -2,449	6 3	0,356 0,092

El análisis de los datos nos muestra que para el contraste en función del tercero de los criterios formulados, las diferencias entre las puntuaciones pretest y postest en la puntuación total resultan significativas al nivel 0,05. Esta significación de las diferencias es compartida por el GE y el GC1; no obstante, hemos de resaltar que para el GE resultan también significativas dos de las dimensiones contempladas en la prueba, una de ellas al nivel 0,01, mientras que el GC1 muestra diferencias estadísticamente significativas para una de las dimensiones. Es de nuestro interés conocer también quien ha conformado este subgrupo de sujetos que puntuaron bajo tanto en ECOS como en la prueba de CI. En el GE, compuesto por 6 niños y una niña, los 7 sujetos recibían AI, de los cuales 4 presentaban discapacidad, 5 recibían también refuerzo de Lengua y dos

presentaban desventaja social. Por lo que respecta al GC1, formado por 3 niños y una niña, hemos hallado en la matriz de datos que 3 de estos sujetos recibían AI, de los cuales 2 presentaban discapacidad y otro recibía también refuerzo de Lengua; sin embargo, el otro alumno no recibía ningún tipo de apoyo ni presentaba discapacidad o privación social. Cuando hemos eliminado a este sujeto de los cálculos realizados anteriormente, el contraste para el GC1 deja de resultar estadísticamente significativo, aunque la diferencia entre las puntuaciones medias sigue siendo muy importante (media pretest: 81,447 (DT = 0,835); media posttest: 91,574 (DT = 5,846); valor $t = -3,031$; sig. 0,094 para 2 gl.).

Nos disponemos a averiguar ahora cómo fue el comportamiento de aquellos sujetos que puntuaron bajo en la prueba de comprensión, pero que sin embargo obtuvieron puntuaciones medias o altas en la prueba de CI. Este perfil se correspondería con los sujetos que han sido descritos como *disléxicos disfonéticos*, como ya recordábamos al analizar los resultados en primaria, presentando especialmente problemas en los procesos de asociación grafemas-fonemas. Hemos realizado los cálculos correspondientes —de forma independiente, como hicimos para primaria, a la valoración que realizó el profesorado sobre este mismo asunto— con relación a los resultados obtenidos por los 5 sujetos que cumplían el perfil. Las puntuaciones observadas para este subgrupo en el GE reflejan ganancias muy significativas entre las fases pretest y posttest. La media pretest para la puntuación total fue de 78,650 (DT = 3,330), mientras que la puntuación posttest fue de 91,382 (DT = 6,892); el valor t obtenido para muestras relacionadas fue de -6,318 y la significación asociada de 0,003 para 4 grados de libertad. En el resto de dimensiones también obtuvo diferencias significativas al nivel 0,01 para *Integración texto-conocimientos* y al nivel 0,05 para *Metaconocimiento estratégico* y *Conocimientos previos*. Las ganancias posttest se produjeron igualmente en el resto de dimensiones, aunque sin llegar a ser significativas. Por otra parte, al igual que hicimos en primaria, hemos querido contrastar los resultados obtenidos por este subgrupo, con la valoración efectuada por el profesorado para estos cinco sujetos sobre la tipología de sus dificultades de aprendizaje. Desgraciadamente, no obtuvimos esta valoración para ninguno de estos alumnos, por lo que no nos ha sido posible efectuar esta comparación.

En cuanto al GC1, tan sólo dos sujetos poseían las condiciones requeridas para este contraste. Estos dos alumnos no obtuvieron ganancias reseñables entre las fases inicial y final del estudio, ni para la puntuación total ni para el resto de las dimensiones contempladas. La puntuación media en el pretest fue de 82,412 (DT = 0,000) y de 83,859 (DT = 8,184) en el posttest; el valor t fue de -0,250 y la significación asociada de 0,844 para 1 grado de libertad. Tampoco en esta ocasión teníamos información sobre la valoración que habían efectuado los docentes para estos alumnos acerca del tipo de sus dificultades de aprendizaje relacionadas con la lectura.

Las DDP en la prueba ECOS

Al igual que hicimos con la prueba de comprensión elaborada para primaria, hemos aplicado la DDP al contraste efectuado con el grupo de sujetos que precisaba de una atención educativa especial, es decir, sujetos con dificultades de aprendizaje, privación social o discapacidad que recibían AI o participaban de programas de DC. Recordemos que el GE estaba formado por 25 sujetos y que la totalidad de éstos se encontraba bajo estas condiciones. Hemos calculado primeramente el número de sujetos

que obtuvieron ganancias o pérdidas en el postest. Los resultados indican que fueron 21 los sujetos que mostraron puntuaciones postest por encima de las iniciales, representando el 84% de la muestra, y 4 los que obtuvieron pérdidas, lo que suponía el 16%. No hubo ningún alumno cuya diferencia de medias fuese cero. Posteriormente hemos otorgado a cada sujeto una puntuación que se correspondía con la magnitud de sus ganancias o pérdidas; hemos vuelto a convertir en porcentaje la suma de los puntos ganadores y perdedores obtenidos, calculando finalmente la puntuación porcentual a favor y en contra de hipótesis que contiene la magnitud de las diferencias. Los resultados quedan expresados a continuación.

	Ganancias	Pérdidas	A favor de hipótesis %	En contra de hipótesis %
Sujetos	21	4	84	16
Puntos	215,56	17,36	92,55	7,45

Podemos observar que la suma de las puntuaciones que expresan ganancias supera ampliamente a la suma de las pérdidas; la puntuación porcentual de la DDP obtenida de esta forma parece subrayar el efecto del tratamiento para este grupo, al tiempo que ofrece información muy valiosa acerca de que los efectos positivos encontrados se han repartido entre la muestra, con un 84% de sujetos que obtiene ganancias postest y más del 90% de la evolución de los puntos favorable a nuestra hipótesis experimental, y no han sido debidos a ganancias de gran intensidad en unos pocos alumnos.

Nos resta establecer la probabilidad con la que pueden ser obtenidos los resultados mostrados. Tal y como expresábamos cuando expusimos los resultados de primaria, la prueba t de Student puede otorgarnos la posibilidad de contrastar las diferencias promedio obtenidas por los sujetos que puntuaron a favor o en contra de la hipótesis. En cuanto a los requisitos para la aplicación de esta prueba, ya comprobamos en análisis anteriores que las puntuaciones obtenidas para la aplicación piloto de la prueba ECOS, donde se recogía la presencia de sujetos con dificultades de aprendizaje de la lectura, presentaban una distribución normal, por lo que la variable *comprensión de textos* estaría distribuida normalmente entre la población; en cualquier caso, hemos procedido a realizar un nuevo contraste mediante la prueba Kolmogorov-Smirnov también para este grupo conformado en su totalidad por sujetos con dificultades de aprendizaje que realizó las pruebas pretest en secundaria. En este caso el valor del estadístico es de 0,109 y la significación asociada de 0,052 para 65 grados de libertad. Al igual que argumentáramos con relación al análisis de la prueba de primaria, hemos optado por aplicar la prueba t en virtud de los resultados del contraste de normalidad para la aplicación piloto y considerando así mismo que la significación asociada al valor t obtenido en este último contraste parece ofrecer suficientes garantías. Los resultados del contraste realizado con la prueba t para muestras independientes se muestran en la siguiente tabla.

Tabla 46. Contraste en función de las puntuaciones DDP para el GE en la prueba ECOS
Prueba t para muestras independientes

	N	Diferencia media (pre — post)	DT	Prueba t	gl.	Sign.
Grupo experimental	21 (G) 4 (P)	10,264 4,340	5,274 4,091	-2,115	23	0,045*

Podemos comprobar que las diferencias obtenidas a favor del grupo de sujetos que experimentó ganancias postratamiento son significativas. También podemos apreciar que la diferencia entre las medias de ambos subgrupos es de casi 6 puntos, lo que supone más del doble de la puntuación de ganancias con respecto a la de pérdidas.

Hemos realizado estos mismos cálculos para el GC1 reducido, es decir, sin tomar en consideración a aquellos sujetos que no precisaban apoyo específico y que no presentaban discapacidad o privación social, ni participaban de programas de diversificación curricular. De los 16 sujetos que conformaban este subgrupo, sólo 3 integraron la columna de pérdidas, de los cuales 2 habían mostrado una diferencia nula entre pretest y postest. La contabilización de estos sujetos en la columna de pérdidas producía finalmente el 18,75% del total; mientras que, por otra parte, 13 alumnos mostraron ganancias, lo que significaba el 81,25%. Gráficamente, las sumas totales, los porcentajes, así como el contraste para evaluar la probabilidad de la DDP se representan a continuación.

	Ganancias	Pérdidas	A favor de hipótesis %	En contra de hipótesis %
Sujetos	13	3	81,25	18,75
Puntos	83,91	3,45	96,05	3,95

*Tabla 47. Contraste en función de las puntuaciones DDP para el GC1 en la prueba ECOS
Prueba t para muestras independientes*

	N	Diferencia media (pre — post)	DT	Prueba t	gl.	Sign.
Grupo	13 (G)	6,454	4,201	-2,129	14	0,052
control	3 (P)	1,148	0,257			

Podemos apreciar que también para este grupo los puntos porcentuales a favor y en contra de la hipótesis contrastan de forma muy notable. Estas diferencias entre las puntuaciones indicadoras de ganancias y pérdidas se reflejan en el contraste realizado mediante la prueba t, contraste que se encuentra al límite de la significatividad estadística.

8.2 Resultados en la escala de ajuste personal-social APSL

La aplicación del dispositivo de evaluación dinámica EDPL supone necesariamente una intervención educativa en el terreno de las actitudes, la motivación, los propios intereses o el modo de afrontar las situaciones de aprendizaje. Este trabajo, basado en procesos de interacción mediada que contemplan la cercanía y la afectividad del evaluador como elementos integrantes del mismo, va encaminado fundamentalmente a provocar la reflexión metacognitiva sobre procesos que están relacionados con el ajuste personal y social del sujeto. Se parte de la base de que los procesos de ajuste personal y social están estrechamente relacionados con el rendimiento escolar, lo que ha sido puesto de manifiesto en numerosas ocasiones (Marchesi, 1999; Mora Merchán y Mora, 2000; Pozo, 1996; Alonso Tapia y Montero, 2001). Aunque lo que se pretende en esencia con la aplicación del dispositivo es, lógicamente, obtener información valiosa acerca de estos procesos y del modo en que pueden optimizarse, contribuyendo a una

mejora global de la adaptación y del rendimiento del individuo, una de nuestras hipótesis contemplaba la posibilidad de que su puesta en práctica (a) incidiera de forma significativa en la modificación de determinadas actitudes relacionadas con el modo de afrontar las situaciones de enseñanza-aprendizaje de la lectura, y (b) modificara algunas ideas, expectativas y representaciones que el sujeto habría contraído y que se verían cuestionadas a la luz del trabajo propuesto. Por otra parte, somos conscientes de que el cambio de actitudes relacionadas con el rendimiento escolar y la adquisición de nuevas representaciones cognitivas o socio-personales, o de nuevas formas de actuar ante determinadas situaciones, es algo que presenta no pocas dificultades en educación (Pozo, 1996; Luque, Molina y Navarro, 2000) y que en todo caso se trata de un proceso educativo complejo y habitualmente de mediana o larga duración. Por tanto no pensábamos de forma simple que con una aplicación limitada como era la que podíamos realizar, se producirían cambios espectaculares en los procesos de ajuste personal-social; máxime cuando, como expresa Mora (1998), muchos de los problemas relacionados con estos procesos de adaptación social y escolar tienen su origen en condiciones y situaciones sociales —aunque también escolares— que quedan fuera del alcance de la influencia de una intervención educativa específica en el marco de una atención educativa especial dentro del contexto escolar.

Por otra parte, de producirse algún cambio en el ámbito de las actitudes, éste debería poder ser evaluado, lo cual nos remite a nuevos problemas, dadas las características esenciales del objeto que nos ocupa. En este sentido, la medición del cambio ha tenido lugar en dos formas: en primer lugar, esta medición se ha efectuado a través del propio proceso de aplicación del dispositivo EDPL, en el que, como se recordará, los aplicadores han recogido información sobre la puesta en práctica por parte del alumnado de habilidades cognitivas, metacognitivas o de ajuste personal-social. Los aplicadores han recogido esta información a través de la interacción mediada llevada a cabo con el alumnado y con base en las actividades propuestas en el dispositivo. Posteriormente, cuando abordemos la evaluación de la aplicación de EDPL, profundizaremos en el contenido de esta información y en las valoraciones cualitativas realizadas por los aplicadores con relación a los procesos de ajuste. La segunda forma de medición del cambio, ésta más susceptible de análisis psicométricos, ha sido la que ha tenido lugar mediante la aplicación de la escala APSL. Con esta escala no podíamos lógicamente acceder a la valoración del proceso de cambio de actitudes, motivación o expectativas en el alumnado, pero sí podíamos, tal como expresábamos cuando analizábamos el contenido de la escala, ofrecer una medida del conocimiento y la percepción que los sujetos tienen sobre sí mismos y sobre sus propias actitudes, motivación, intereses, expectativas o modo de afrontar las situaciones problemáticas relacionadas con el aprendizaje. En este sentido, el instrumento podría ser considerado un modo de evaluación del conocimiento metacognitivo con relación al desarrollo socio-personal. La metacognición constituiría un pilar fundamental del desarrollo social y personal del sujeto, y las alteraciones en los procesos metacognitivos podrían estar en la base de las dificultades de aprendizaje, así como de los trastornos relacionados con el ajuste personal-social (Mora, 1994, 1997; Mora Merchán y Mora, 2000).

De este modo, los resultados de la escala APSL deberían aportarnos información sobre el conocimiento y la percepción que los sujetos tienen sobre sí mismos, sus dificultades, su motivación, sus intereses y sus expectativas. Los contrastes pretest-postest deberían aportarnos así mismo información sobre el posible grado de cambio que la aplicación del dispositivo EDPL ha originado en el grupo experimental, y si este cambio es

diferente al que habría tenido lugar de no aplicarse el tratamiento, opción ésta representada por el grupo control. Tal y como apuntábamos anteriormente, la aplicación del dispositivo debería originar cambios a favor de la hipótesis experimental en el grupo de tratamiento, aún cuando éstos, debido a los condicionantes que exponíamos más arriba, así como al contenido del objeto evaluado, no fueran muy significativos. Sin embargo, en lo que respecta al GC1, se esperaba que estos cambios fueran inexistentes o al menos inestables a lo largo de la muestra, es decir, que pudieran ser explicados por la acción del azar. Nos disponemos por tanto a continuación a revisar y exponer los resultados obtenidos con la aplicación pretest y posttest de la escala APSL.

Realizaremos diversos contrastes en función de criterios que nos permitan extraer el máximo de información acerca de cómo, por qué y con quien ha funcionado el dispositivo. Serán utilizados básicamente los tres criterios que ya formulamos para el análisis de las pruebas de comprensión, es decir: (a) designación del profesorado y de los orientadores colaboradores acerca de la condición de sujetos con dificultades de aprendizaje de la lectura; (b) designación del profesorado más una puntuación baja obtenida en la fase pretest de la escala; y (c) designación del profesorado más una puntuación baja tanto en la escala de actitudes como en la prueba de CI. En este sentido, tal y como ya expresábamos con anterioridad para las pruebas de comprensión, dado que las puntuaciones directas obtenidas en la escala han sufrido idéntico proceso de transformación que en éstas, hemos considerado igualmente que una puntuación baja era aquella que se encontraba por debajo de una desviación típica de la media, es decir, una puntuación de 84 o inferior. En principio, estos contrastes se realizarán tomando en consideración a la totalidad del alumnado participante en el estudio, ya pertenezcan a la etapa de primaria o de secundaria, ya que, como se recordará, el contenido de la escala no tuvo variación significativa para ambas poblaciones y la evaluación se previó conjunta; no obstante, abordaremos el contraste igualmente por separado para evaluar el posible impacto del trabajo realizado en función de esta condición. Por otra parte, en el capítulo dedicado al análisis de la composición de la muestra, hacíamos referencia a la equivalencia inicial de los grupos experimental y control, por lo que nos disponemos ahora a acometer la evaluación de las posibles diferencias postratamiento.

(a) Criterio de designación previa del profesorado sobre las dificultades de aprendizaje

El primer análisis que efectuamos es para el grupo completo de alumnos y alumnas que fueron catalogados por el profesorado como sujetos con dificultades de aprendizaje, tanto en la etapa de primaria como en secundaria. Exponemos en la tabla siguiente los resultados de puntuaciones medias obtenidas en las fases inicial y final del estudio, así como los valores del contraste realizado, tanto para la puntuación global obtenida como para las distintas dimensiones contempladas en la escala.

Puede comprobarse que en función de este primer criterio, los resultados no muestran ninguna significatividad para la puntuación global; las puntuaciones pretest y posttest se mantienen prácticamente iguales, aunque con pequeños incrementos postratamiento en el GE, especialmente para algunas de las dimensiones contempladas, y no varían sustancialmente para el GC1, aunque experimentan una ligera pérdida. Hay que destacar en cuanto a las dimensiones, las diferencias favorables al posttest halladas para la D9 en el GE, acercándose a la significatividad estadística.

Dimensiones de la escala APSL

- D1 Conocimiento de las propias limitaciones cognitivas actuales, potencialidades y posibilidades de aprendizaje, dificultades y vías de solución.
- D2 Conocimiento de las propias expectativas, autoconcepto académico y estilo atribucional.
- D3 Conocimiento de la propia actitud ante la lectura: emociones, sentimiento de rechazo vs satisfacción y gusto por la lectura, tolerancia a la frustración, persistencia ante la dificultad, autonomía vs dependencia.
- D4 Conocimiento de los propios intereses y motivaciones.
- D5 Percepción del apoyo que ofrece el contexto escolar y familiar, y de las condiciones en las que se desenvuelve.
- D6 Percepción de la relevancia del tema: la lectura, para que sirve, cual es su importancia.
- D7 Conocimiento de las propias estrategias de relación con los demás en tareas de lectura.
- D8 Conocimiento de las propias estrategias de control y regulación consciente del proceso de lectura (planificación, supervisión y autoevaluación).
- D9 Reconocimiento de la capacidad de generalización de estos conocimientos o estrategias.

Tabla 48. Comparación pretest-postest para la escala de ajuste personal-social APSL en función del criterio de designación previa del profesorado

Gr.	Dimensiones	N	Pretest	DT	Postest	DT	T	gl	Sign.
GE	Puntuación global	54	97,823	14,054	97,928	14,753	-0,067	53	0,947
GC		56	95,635	15,877	93,412	16,069	1,106	55	0,273
GE	D1 Conocimiento de las propias limitaciones...	54	94,283	15,885	97,033	15,161	-1,398	53	0,168
GC		56	95,613	14,513	96,336	15,592	-0,358	55	0,721
GE	D2 Conocimiento de las propias expectativas...	54	98,408	15,549	97,100	16,800	0,564	53	0,575
GC		56	94,433	15,700	93,592	14,495	0,457	55	0,650
GE	D3 Conocimiento de la propia actitud...	54	95,918	16,853	96,687	13,738	-0,345	53	0,731
GC		56	96,726	14,843	93,865	16,639	1,199	55	0,236
GE	D4 Conocimiento de los propios intereses...	54	99,630	13,979	100,361	15,329	-0,386	53	0,701
GC		56	97,626	15,490	96,215	17,140	0,715	55	0,478
GE	D5 Percepción del apoyo del contexto...	54	104,221	15,377	99,453	16,691	1,997	53	0,051
GC		56	99,619	15,680	97,078	18,270	0,941	55	0,351
GE	D6 Percepción de la relevancia del tema...	54	100,008	14,252	97,124	18,616	1,032	53	0,307
GC		56	100,142	14,086	94,579	19,083	2,629	55	0,011(-)
GE	D7 Conoc. de propias estrategias de relación...	54	102,225	17,196	103,125	15,891	-0,327	53	0,745
GC		56	99,085	15,290	96,771	17,140	1,000	55	0,322
GE	D8 Metaconocimiento estratégico de lectura	54	97,401	13,317	98,662	15,080	-0,580	53	0,564
GC		56	95,499	16,381	94,588	16,373	0,335	55	0,739
GE	D9 Reconocimiento de cap. de generalización...	54	95,867	16,528	101,485	15,074	-1,917	53	0,061
GC		56	97,372	17,068	100,467	15,730	-1,306	55	0,197
GE	Dimensiones 8 y 9	54	96,259	12,940	99,331	15,125	-1,404	53	0,166
GC		56	95,097	15,568	95,367	16,944	-0,098	55	0,922

Hemos procedido a realizar estos contrastes para el grupo reducido, es decir, aquel conformado por los sujetos con DA que reciben AI, que presentan discapacidad, privación social o participan de programas de DC en secundaria. Al igual que ya ocurriera cuando analizábamos los resultados de las pruebas de comprensión, el GE reduce poco sus efectivos; tan sólo en dos alumnos. La puntuación global no varía y se mantiene en pretest y postest en unos 97 puntos. Para el resto de dimensiones de la escala tampoco hay variación sustancial. En cambio para el GC1 el número de sujetos, como viene siendo habitual en los contrastes que hemos realizado, desciende de forma significativa, hasta 30. Las diferencias siguen sin ser significativas para la puntuación

global, con una ligera pérdida en la fase posttest; la media pretest se sitúa en 93,464 (DT = 15,325) y la posttest en 92,097 (DT = 15,627), con un valor t de 0,667 y una significación asociada de 0,510 para 29 grados de libertad. En cuanto a las dimensiones, los resultados muestran diferencias significativas a favor de la puntuación posttest al nivel 0,05 para la D1, mientras que para la D6 las diferencias resultan significativas, también al nivel 0,05, pero a favor de las puntuaciones pretest.

Hemos comprobado cómo ha sido el comportamiento de los sujetos en la escala en función de la etapa. Preveíamos que pudieran existir diferencias notables, no sólo en cuanto a los niveles de ajuste personal-social con relación a la edad y el nivel educativo de los sujetos, sino también en cuanto al modo de afrontar una tarea como esta, en la que, en definitiva, se pide al alumnado que realice una reflexión acerca de asuntos concernientes a las propias actitudes frente a lo escolar, a su motivación, sus intereses o su percepción del apoyo que recibe, tanto dentro como fuera del contexto escolar. Ante estos elementos, es esperable en algún modo que el comportamiento y la actitud de los sujetos sea distinto en función de la edad, lo que viene a querer decir, en función de las experiencias acumuladas y de cómo han sido vividas y asimiladas esas experiencias. Los resultados de estos contrastes para el GE los mostramos en la siguiente tabla.

Tabla 49. Comparación pretest-posttest en primaria y secundaria para la escala de ajuste personal-social APSL en función del criterio de designación previa del profesorado para el GE

Gr.	Dimensión	Etapa	N	Pretest	DT	Posttest	DT	T	gl	Sign.
GE	Puntuación global	EP	28	98,287	13,810	98,136	13,570	0,073	27	0,942
		ESO	26	97,324	14,570	97,704	16,200	-0,157	25	0,876
GE	D1	EP	28	93,202	15,156	96,819	15,006	-1,191	27	0,244
		ESO	26	95,446	16,856	97,264	15,619	-0,727	25	0,474
GE	D2	EP	28	98,356	14,956	97,795	16,310	0,193	27	0,849
		ESO	26	98,464	16,461	96,351	17,604	0,570	25	0,574
GE	D3	EP	28	99,057	17,268	96,726	14,357	0,786	27	0,438
		ESO	26	92,537	16,037	96,645	13,322	-1,250	25	0,223
GE	D4	EP	28	102,362	12,278	105,183	13,942	-1,039	27	0,308
		ESO	26	96,688	15,301	95,169	15,302	0,582	25	0,566
GE	D5	EP	28	102,643	14,087	99,014	15,930	1,161	27	0,256
		ESO	26	105,919	16,768	99,926	17,780	1,625	25	0,117
GE	D6	EP	28	98,195	15,699	94,301	20,915	0,864	27	0,395
		ESO	26	101,961	12,521	100,163	15,612	0,550	25	0,587
GE	D7	EP	28	98,796	16,997	101,689	16,106	-0,739	27	0,466
		ESO	26	105,918	16,957	104,672	15,824	0,319	25	0,753
GE	D8	EP	28	98,538	11,423	97,019	16,512	0,452	27	0,655
		ESO	26	96,177	15,233	100,431	13,468	-1,612	25	0,119
GE	D9	EP	28	93,889	16,132	96,211	15,524	-0,486	27	0,631
		ESO	26	97,997	16,998	107,165	12,520	-2,848	25	0,009**
GE	D8 y D9	EP	28	96,578	10,404	96,040	15,941	0,154	27	0,878
		ESO	26	95,916	15,422	102,875	13,615	-2,900	25	0,008**

El contraste realizado dividiendo las muestras de primaria y secundaria, si bien no ha aportado cambios a las diferencias establecidas entre las fases inicial y final con respecto a la puntuación global, nos ha permitido observar algunas diferencias

estadísticamente significativas para algunas de las dimensiones contempladas en la escala. Estas diferencias son especialmente importantes para el alumnado de secundaria en las dimensiones relacionadas con el metaconocimiento estratégico sobre la lectura (D8) y con el reconocimiento de la capacidad de generalizar esas estrategias (D9), llegando a ser muy significativas para ésta última, así como para la dimensión que aún los aspectos citados (D8 y 9). Estos resultados nos ayudan a matizar los resultados globales y nos comienzan a ofrecer algunas pistas sobre el perfil del alumnado que se podría haber beneficiado en mayor medida del trabajo relacionado con los procesos de ajuste personal-social, así como sobre qué dimensiones de las contempladas en la escala ha tenido un mayor efecto este trabajo.

Por lo que respecta al GC1, la puntuación global para la etapa de primaria indica una ligera pérdida postest, con una media inicial de 97,732 (DT = 19,643) y una postest de 96,105 (DT = 20,608); un valor *t* de 0,396 y una significación asociada de 0,696 para 19 grados de libertad. En cuanto a los contrastes para las distintas dimensiones, los resultados no arrojan diferencias reseñables a favor del postest, pero sí diferencias significativas a favor de las puntuaciones iniciales al nivel 0,05 para la dimensión 6. En secundaria, las diferencias para la puntuación global son negativas, es decir, a favor del pretest, con una media inicial de 94,470 (DT = 13,519) y una postest de 91,916 (DT = 12,991); con un valor *t* de 1,169 y una significación de 0,250 para 35 gl. Los contrastes para las dimensiones muestran ausencia de diferencias reseñables, siendo la mayoría de ellas de orden negativo.

Hemos contrastado las diferencias de medias pretest-postest entre el GE y el GC1 con el objetivo de comprobar la existencia de contrastes entre diferencias promedio que pudieran resultar significativos. En primer lugar hemos efectuado estos cálculos para el grupo completo de sujetos, no hallando ningún contraste entre las diferencias de medias que fuese significativo. Para la puntuación global el valor de la diferencia entre las puntuaciones que expresan las diferencias de medias fue de 2,327 puntos favorables al GE, con un valor *t*, asumiendo la igualdad de varianzas, de -0,908 y una significación asociada de 0,336 para 108 grados de libertad. Cuando hemos realizado el contraste para el alumnado de primaria tampoco hallamos diferencias que resultaran estadísticamente significativas, siendo el valor de la diferencia entre las medias de 1,475 a favor del GE y el valor *t* para la puntuación global de -0,348, con una significación de 0,730 para 46 gl. Por otra parte, para la etapa de secundaria, el valor de la diferencia entre las puntuaciones diferencia fue de 2,934 a favor del GE, con un valor *t* de -0,891 y una significación de 0,377 para 60 gl; en cambio, los contrastes para las distintas dimensiones consideradas en la escala muestran aquí algunas diferencias importantes. En concreto, la variación entre las puntuaciones diferencia para la dimensión D8 y 9 resulta significativa a favor del GE. El valor de esta diferencia es de 10,100 puntos, con un valor *t* de -2,458 y una significación asociada de 0,017. Para las dimensiones D8, con un valor de diferencia de más de 8 puntos, y para la D9, con un valor cercano a los 8 puntos, el contraste se acerca a la significatividad estadística (sig.0,065 para D8 y sig. 0,086 para D9).

Por otra parte, también nos parecía interesante contrastar los resultados de chicos y chicas en virtud de las posibles diferencias que se pudieran encontrar en la realización de una escala como APSL. Numerosas investigaciones han puesto de manifiesto la existencia de diferencias a favor de las chicas en cuanto a rendimiento en lectura —si bien estas diferencias parecen neutralizarse una vez superados los primeros cursos de la

educación secundaria—, así como también en cuanto a la adaptación social y escolar (Mora, 1998), lo que se traduce en la actitud, la motivación y los intereses que se manifiestan en el contexto escolar. Estas diferencias, como hemos apuntado, parecen ser mayores en la etapa de primaria y decrecen sobre todo al final de la etapa de secundaria. Hemos realizado estos contrastes sirviéndonos de las puntuaciones diferencia, es decir, aquellas que expresan la diferencia entre la puntuación obtenida en el pretest y la que se obtuvo en el postest.

Cuando hemos comparado las puntuaciones de chicos y chicas tomando en consideración a la totalidad del alumnado con dificultades de aprendizaje que participó en el estudio, ya fuera formando parte del GE o del GC1, no hemos encontrado diferencias significativas, con un valor de la diferencia para la puntuación global de 0,363 favorable a las niñas y un valor t de $-0,136$, con una significación asociada de 0,892 para 108 gl. Este grupo de comparación lo formaban 40 niñas y 70 niños. En lo referente a la submuestra de primaria, los resultados muestran igualmente ausencia de diferencias estadísticamente reseñables para la puntuación global; en cambio, sí es destacable que el valor de la diferencia sea favorable a los niños (3,838 puntos), con un valor t de 0,884 y una significación de 0,381 para 46 gl. Este subgrupo estaba formado por 17 niñas y 31 niños. En este mismo contraste y para la dimensión D2, las diferencias se acentúan y llegan a ser muy significativas, con un valor de diferencia de 11,851 puntos favorables a los chicos, un valor t de 2,712 y una significación de 0,009 para 46 gl. Por su parte, la submuestra de secundaria tampoco muestra diferencias significativas para la puntuación global. Aquí el valor de diferencia es de 3,570, pero favorable esta vez a las chicas; el valor t de este contraste se sitúa en $-1,064$ y la significación que se asocia a este valor, de 0,292 para 60 gl. Para la dimensión D5 los resultados son casi significativos, con un valor de diferencia de más de 8 puntos favorable a las chicas y una significación asociada de 0,063.

En general podemos concluir que para la puntuación global de la escala no se han encontrado diferencias significativas entre chicos y chicas, ni en la muestra completa ni cuando los contrastes han sido efectuados en función de la etapa educativa. Para la submuestra de primaria, las diferencias halladas han sido ligeramente favorables a los chicos, diferencias que llegaban a ser significativas para una de las dimensiones consideradas (D2); mientras que en secundaria, las diferencias eran más favorables a las chicas, acercándose a la significatividad estadística el contraste para la dimensión D5.

Las anteriores comparaciones nos hablan del comportamiento de chicos y chicas sin tomar en consideración su participación del GE o del GC1; sin embargo, resulta de nuestro interés conocer si ha existido un comportamiento diferencial de los sujetos que participaban del GE en función del sexo, ya que esto nos hablaría de una respuesta diferenciada al tratamiento cuando el objeto de evaluación es el ajuste personal-social. En este sentido, hemos contrastado las puntuaciones diferencia para el GE completo, no hallando diferencias reseñables para la puntuación global, con valores muy similares en las diferencias promedio. En cambio, sí se mostraron significativas estadísticamente a favor de las chicas las diferencias con relación a la dimensión D5, con un valor de diferencia de 12,415 puntos, un valor t de $-2,650$ y una significación de 0,011 para 52 grados de libertad. Este subgrupo está formado por 20 niñas y 34 niños. Cuando dividimos el GE en función de la etapa, los resultados muestran ausencia de diferencias significativas para los contrastes de la puntuación global; sin embargo, para primaria, indican la significatividad de las comparaciones en la dimensión D2, con un valor de

diferencia de más de 17 puntos, y en la D3, con un valor de diferencia superior a los 15 puntos, ambas favorables a los chicos. En secundaria tampoco se hallan diferencias significativas para la puntuación global, aunque en este caso el valor de la diferencia es de casi 5 puntos a favor de las chicas. Donde sí se encuentran diferencias significativas es en el contraste para la dimensión D5, con un valor de diferencia de más de 15 puntos favorables a las chicas.

Contrastes en función de los distintos subgrupos considerados en la muestra

Al igual que hemos venido realizando para las pruebas de comprensión, nos disponemos ahora a efectuar los contrastes en función de los distintos subgrupos que hemos considerado en nuestro estudio. El análisis de estos datos nos interesa especialmente, ya que, paradójicamente, amplía sobremanera los horizontes de nuestras conclusiones. En efecto, la aplicación del dispositivo EDPL, y, en este caso, en lo relativo a los procesos de ajuste, podría haber tenido un efecto diferencial en función del trabajo realizado con sujetos que presentan distintas características relativas al grado de sus dificultades o a sus condiciones de apoyo. Con estos análisis queremos acercarnos a este tipo de análisis que, sin duda, quedará completado cuando abordemos la evaluación complementaria del dispositivo y de los resultados en el próximo capítulo. En la siguiente tabla mostramos los resultados obtenidos para los grupos experimental y control.

Tabla 50. Comparaciones pretest-postest en la escala APSL (puntuación global) para los distintos subgrupos contemplados en la muestra (DA = dificultades de aprendizaje; DIS = discapacidad; AI = recibe apoyo a la integración; DEP = privación socio-cultural; RL = recibe refuerzo de Lengua; NA. = no recibe apoyo; AC = apoyo a ciclo; AF = agrupamientos flexibles; DC = diversificación curricular)

Prueba t para muestras relacionadas

Grupos	N	Pretest	DT	Postest	DT	valor t	gl	significación
GE DA	29	96,232	14,351	98,280	14,582	-0,862	28	0,396
GC1 DA	33	95,985	15,613	93,156	16,252	0,959	32	0,345
GE DIS	13	97,759	12,401	92,862	14,997	1,916	12	0,080
GC1 DIS	16	97,378	11,468	95,079	11,777	1,220	15	0,241
GE AI	37	95,103	13,201	95,830	14,096	-0,343	36	0,734
GC1 AI	26	92,645	13,358	92,155	12,721	0,281	25	0,781
GE DEP	12	101,740	15,371	102,565	14,423	-0,297	11	0,772
GC1 DEP	7	90,002	25,221	90,810	24,411	-0,109	6	0,917
GE NA, AC, RL	8	104,451	14,165	105,158	12,215	-0,231	7	0,824
GC NA, AC, RL	26	98,140	16,430	94,930	16,743	0,876	25	0,390

Para el primero de los contrastes, en el que los sujetos de ambos grupos fueron catalogados con DA por sus profesores u orientadores, en ausencia de discapacidad o privación socio-cultural, podemos observar con relación a la puntuación global que las puntuaciones expresan ganancias postest para el GE, aunque éstas no llegan a ser significativas. Por otra parte, aunque no aparecen en esta tabla, hemos extraído los datos correspondientes al contraste para las distintas dimensiones contempladas. En este sentido, la mayoría de las dimensiones expresan diferencias favorables al postest,

especialmente aquellas que hacen referencia al metaconocimiento estratégico (D8) y al reconocimiento de la capacidad de generalización de las estrategias (D9), con valores de significación cercanos a la significatividad estadística. Finalmente, es la dimensión compuesta (D8 y 9) la que obtiene diferencias significativas, con una diferencia de medias de 7,019 puntos favorables al postest, un valor t de -2,486 y una significación asociada de 0,019 para 28 grados de libertad. En lo referente al contraste para el GC1, además de las pérdidas que podemos observar para la puntuación global, es destacable que la mayoría de las comparaciones en las dimensiones contempladas resultan favorables a la puntuación inicial y, por tanto, indican pérdidas que, en cualquier caso, no llegan a ser significativas.

Hemos realizado el contraste de diferencias medias entre los grupos experimental y control de este subgrupo. Los resultados expresan valores de diferencia pretest-postest favorables al GE para la mayoría de las dimensiones, así como para la puntuación global de la escala. En ésta, el valor es de 4,877 puntos, la t observada es de -1,265 y la significación asociada de 0,211 para 60 gl., asumiendo igualdad de las varianzas. En la dimensión D1 las diferencias llegan a ser significativas, con un valor de la diferencia de 8,161 puntos, un valor t de -2,083 y una significación de 0,042 para 60 gl., también asumiendo la igualdad de varianzas.

En el segundo de los contrastes realizados, el que hace referencia a los sujetos con *discapacidad*, los datos muestran un decremento notable de la puntuación global en el postest para el GE; en los contrastes para las distintas dimensiones, la mayoría son favorables a la puntuación pretest, expresando algunas de ellas importantes pérdidas, aunque ninguna de ellas llega a la significatividad estadística. Por su parte, el GC1 también obtiene pérdidas postest, aunque de menor intensidad. En cuanto a los resultados de este grupo en las distintas dimensiones, ningún contraste resulta significativo, indicando igualmente la mayoría un decremento de las puntuaciones finales. El contraste de diferencias medias entre GE y GC1, realizado posteriormente, no arroja ningún valor significativo ni destacable.

El contraste en función de la condición de *desventaja socio-cultural* muestra un ligero incremento postest de las puntuaciones del GE, pero a todas luces insuficiente para el rechazo de la hipótesis de igualdad de medias entre las mediciones. En cuanto a las dimensiones en este grupo, ninguno de los contrastes resulta significativo, mostrando la mayoría pequeñas ganancias postratamiento. El GC1, partiendo de una puntuación inicial notablemente inferior a la del GE, sin que estas diferencias tengan significación estadística ($t = 1,270$; sig. 0,221 para 17 gl), muestra igualmente ausencia de diferencias significativas para la puntuación global. Los contrastes para las dimensiones muestran en su mayoría ligeras ganancias postest, aunque sin llegar a la significatividad estadística. El contraste de diferencias promedio para este subgrupo no muestra diferencias reseñables para la puntuación global ni tampoco para las dimensiones contempladas, a pesar de que para algunas de ellas existen valores de diferencia notables entre los grupos. Algunos de estos valores son favorables al GE y otros al GC1.

Para el contraste en función de la condición *recibe apoyo a la integración*, los resultados para la puntuación global en ambos grupos muestran una similitud cuanto menos sorprendente. En cuanto a las dimensiones de la escala para el GE, tan sólo la D9 se acerca a la significatividad estadística; el resto de contrastes expresa ganancias o pérdidas ligeras. En el GC1, el contraste para las dimensiones arroja valores de

ganancias o pérdidas moderadas, excepto para la dimensión D6, donde se producen pérdidas que llegan a ser muy significativas, con un valor de diferencia de 8,387 puntos, un valor t de 2,776 y una significación de 0,010 para 25 grados de libertad. El contraste de diferencias de medias no arrojó ningún valor significativo, con valores de diferencias moderados, excepto también para la dimensión D6, donde el valor fue superior a los 7 puntos favorables al GE, a pesar de lo cual, el contraste no resultó estadísticamente significativo. El valor t observado fue de -1,446 y la significación asociada de 0,153 para 61 grados de libertad.

Nos resta aludir al subgrupo de sujetos formado por aquellos que, bien no recibían ningún tipo de apoyo, o bien recibían apoyo a ciclo en primaria o refuerzo de Lengua en secundaria, y que no presentaban discapacidad o deprivación social, es decir, supuestamente, aquellos sujetos con un nivel menor de dificultad. Al igual que hicimos cuando abordamos el análisis de la prueba de comprensión en secundaria, hemos incluido en este subgrupo a los alumnos de 4º de ESO que participaban de programas de DC, salvo a los que presentaran discapacidad o deprivación social, que habrían sido ya incluidos en otros subgrupos. Como podemos comprobar, el número de sujetos que se hallaban en estas condiciones es muy dispar entre el GE y el GC1. Tan sólo 8 alumnos, de los cuales 6 participaban de DC, cumplían este criterio en el GE, mientras que 26 lo hacían para el GC1. Las puntuaciones obtenidas por los sujetos que recibieron el tratamiento muestran un ligero incremento de la puntuación posttest para la puntuación global, aunque lejos de la significatividad estadística. Por otra parte, la mayoría de las dimensiones expresan diferencias favorables al posttest, con valores de diferencia por encima de los 8 puntos. Los contrastes para las dimensiones D8 y D9 están cercanos a la significatividad. En cuanto al GC1, el contraste para la puntuación global muestra pérdidas, aunque no significativas, al igual que la mayoría de las comparaciones para las distintas dimensiones. Por último, el contraste de diferencias medias arroja valores de diferencia importantes, que llegan a superar los 13 y los 11 puntos favorables al GE para las dimensiones D7 y D9 respectivamente; estos contrastes tienen una significación asociada al valor t cercana al nivel 0,05.

Contrastes en función de los grupos de aplicación del dispositivo

Al igual que argumentábamos para las pruebas de comprensión la conveniencia de realizar comparaciones en función de los distintos subgrupos del GE a quien un mismo profesional había aplicado el dispositivo EDPL, consideramos de gran interés realizar estos contrastes también para los resultados obtenidos en la escala APSL. Somos conscientes de que al realizar estos contrastes, el tamaño de las submuestras resultantes es extremadamente reducido, lo que sin duda tiene una influencia notable en los resultados de las pruebas estadísticas aplicadas. A pesar de esta circunstancia, creemos que puede ser útil y reveladora la exploración de los datos comparativos entre las fases inicial y final para cada uno de los subgrupos, así como entre los resultados obtenidos en función del aplicador. Lógicamente, uno de los motivos que hacen de esta tarea algo relevante, es el de contrastar a su vez los resultados para las distintas pruebas aplicadas. En efecto, si los sujetos de un mismo aplicador obtuvieran ganancias en distintas pruebas que, aunque relacionadas, evalúan aspectos y procesos distintos con relación a la lectura, podríamos afirmar con más certeza la valía y la validez de estas ganancias, que habrían tenido lugar tanto en procesos relacionados con el rendimiento en lectura, como en aquellos más relacionados con las actitudes. Por otra parte, también incrementaría la validez y la fiabilidad del cambio al ofrecer resultados en parte

convergentes obtenidos de distintas pruebas. Nos disponemos ahora a realizar estas comparaciones en función de los aplicadores que han prestado su colaboración en el estudio.

Tabla 51. Comparaciones pretest-postest para la escala APSL por grupos de aplicadores
Estadísticos de muestras relacionadas

Grupos	Dimensiones		Media	N	DT	t	gl	Sign.
Grupo A CEIP Andalucía	Puntuación total	Pretest	108,223	6	8,739	1,224	5	0,276
		Postest	103,744		15,235			
	D1 Conocimiento de las propias limitaciones...	Pretest	98,908	6	14,425	-1,185	5	0,289
		Postest	104,534		14,160			
	D2 Conocimiento de las propias expectativas...	Pretest	107,416	6	16,207	1,190	5	0,287
		Postest	98,262		14,396			
	D3 Conocimiento de la propia actitud...	Pretest	108,663	6	8,733	1,695	5	0,151
		Postest	99,763		18,255			
	D4 Conocimiento de los propios intereses...	Pretest	110,288	6	9,214	0,349	5	0,741
		Postest	109,347		13,071			
	D5 Percepción del apoyo del contexto...	Pretest	108,612	6	11,064	0,200	5	0,849
		Postest	107,483		8,298			
	D6 Percepción de la relevancia del tema...	Pretest	106,354	6	8,043	1,659	5	0,158
Postest		88,181	25,595					
D7 Conoc. de propias estrategias de relación...	Pretest	106,126	6	18,941	0,000	5	1,000	
	Postest	106,126		12,197				
D8 Metaconocimiento estratégico de lectura	Pretest	98,031	6	9,945	-0,542	5	0,611	
	Postest	99,449		12,031				
D9 Reconocimiento de cap. de generalización...	Pretest	98,275	6	16,312	0,000	5	1,000	
	Postest	98,275		16,312				
Dimensiones 8 y 9	Pretest	97,655	6	6,743	-0,349	5	0,741	
	Postest	98,912		12,985				

Grupos	Dimensiones		Media	N	DT	t	gl	Sign.
Grupo B CEIP Antonio Machado	Puntuación total	Pretest	95,492	6	11,849	2,122	5	0,087
		Postest	91,012		11,102			
	D1 Conocimiento de las propias limitaciones...	Pretest	97,783	6	6,038	1,274	5	0,259
		Postest	92,157		9,936			
	D2 Conocimiento de las propias expectativas...	Pretest	93,032	6	14,679	-0,237	5	0,822
		Postest	94,339		20,259			
	D3 Conocimiento de la propia actitud...	Pretest	93,830	6	9,689	-0,725	5	0,501
		Postest	97,786		10,613			
	D4 Conocimiento de los propios intereses...	Pretest	102,765	6	12,361	0,395	5	0,709
		Postest	100,884		17,360			
	D5 Percepción del apoyo del contexto...	Pretest	99,578	6	19,481	1,168	5	0,296
		Postest	92,803		12,615			
	D6 Percepción de la relevancia del tema...	Pretest	90,778	6	16,087	0,591	5	0,580
Postest		82,989	22,928					
D7 Conoc. de propias estrategias de relación...	Pretest	92,624	6	8,367	-1,936	5	0,111	
	Postest	108,826		13,229				
D8 Metaconocimiento estratégico de lectura	Pretest	99,449	6	13,179	2,210	5	0,078	
	Postest	81,016		13,629				
D9 Reconocimiento de cap. de generalización...	Pretest	94,663	6	13,705	0,415	5	0,695	
	Postest	91,051		16,312				
Dimensiones 8 y 9	Pretest	97,655	6	11,679	2,150	5	0,084	
	Postest	80,064		13,202				

Tabla 51. Comparaciones pretest-postest para la escala APSL por grupos de aplicadores
(Continuación) Estadísticos de muestras relacionadas

Grupos	Dimensiones		Media	N	DT	t	gl	Sign.
Grupo C CEIP El Ruedo	Puntuación total	Pretest	84,883	3	4,547 17,111	-0,750	2	0,531
		Postest	93,370					
	D1 Conocimiento de las propias limitaciones...	Pretest	75,281	3	3,897 14,052	-1,732	2	0,225
		Postest	88,782					
	D2 Conocimiento de las propias expectativas...	Pretest	89,108	3	7,846 16,333	-0,400	2	0,728
		Postest	94,339					
	D3 Conocimiento de la propia actitud...	Pretest	81,964	3	27,404 20,553	-0,277	2	0,808
		Postest	85,919					
	D4 Conocimiento de los propios intereses...	Pretest	91,480	3	9,772 5,642	-1,964	2	0,188
		Postest	108,407					
	D5 Percepción del apoyo del contexto...	Pretest	99,578	3	10,349 25,651	0,000	2	1,000
		Postest	99,578					
	D6 Percepción de la relevancia del tema...	Pretest	101,162	3	8,993 23,793	1,000	2	0,423
Postest		90,778						
D7 Conoc. de propias estrategias de relación...	Pretest	98,025	3	24,750 24,750	0,378	2	0,742	
	Postest	92,624						
D8 Metaconocimiento estratégico de lectura	Pretest	93,778	3	4,911 21,410	-0,277	2	0,808	
	Postest	96,613						
D9 Reconocimiento de cap. de generalización...	Pretest	94,663	3	21,669 25,021	-0,277	2	0,808	
	Postest	101,886						
Dimensiones 8 y 9	Pretest	92,629	3	8,705 27,182	-0,286	2	0,802	
	Postest	97,655						

Grupos	Dimensiones		Media	N	DT	t	gl	Sign.
Grupo D CEIP El Ruedo	Puntuación total	Pretest	74,745	2	1,000 6,001	-3,800	1	0,164
		Postest	88,183					
	D1 Conocimiento de las propias limitaciones...	Pretest	77,531	2	9,546 19,093	-4,000	1	0,156
		Postest	104,534					
	D2 Conocimiento de las propias expectativas...	Pretest	89,108	2	0,000 5,548	1,000	1	0,500
		Postest	85,185					
	D3 Conocimiento de la propia actitud...	Pretest	79,986	2	8,391 4,195	-0,333	1	0,795
		Postest	82,952					
	D4 Conocimiento de los propios intereses...	Pretest	85,837	2	7,979 7,979	-1,500	1	0,374
		Postest	102,765					
	D5 Percepción del apoyo del contexto...	Pretest	87,156	2	14,373 14,373	No se puede calcular t		
		Postest	80,381					
	D6 Percepción de la relevancia del tema...	Pretest	80,393	2	22,028 22,028	No se puede calcular t		
Postest		80,393						
D7 Conoc. de propias estrategias de relación...	Pretest	79,121	2	11,457 11,457	No se puede calcular t			
	Postest	111,527						
D8 Metaconocimiento estratégico de lectura	Pretest	86,688	2	18,047 18,047	0,000	1	1,000	
	Postest	86,688						
D9 Reconocimiento de cap. de generalización...	Pretest	72,994	2	0,000 15,322	-3,000	1	0,205	
	Postest	105,498						
Dimensiones 8 y 9	Pretest	78,808	2	15,992 10,661	-0,600	1	0,656	
	Postest	90,116						

Tabla 51. Comparaciones pretest-postest para la escala APSL por grupos de aplicadores
(Continuación) Estadísticos de muestras relacionadas

Grupos	Dimensiones		Media	N	DT	t	gl	Sign.
Grupo E CEIP Lope de Rueda	Puntuación total	Pretest	110,345	3	4,321 12,196	-0,311	2	0,785
		Postest	111,760					
	D1 Conocimiento de las propias limitaciones...	Pretest	93,282	3	15,589 20,623	-3,500	2	0,073
		Postest	109,034					
	D2 Conocimiento de las propias expectativas...	Pretest	115,263	3	4,530 4,530	-1,000	2	0,423
		Postest	117,878					
	D3 Conocimiento de la propia actitud...	Pretest	105,697	3	9,063 3,425	0,555	2	0,635
		Postest	101,741					
	D4 Conocimiento de los propios intereses...	Pretest	106,526	3	11,745 11,745	-1,732	2	0,225
		Postest	112,169					
	D5 Percepción del apoyo del contexto...	Pretest	117,646	3	6,775 14,104	0,945	2	0,444
		Postest	106,354					
	D6 Percepción de la relevancia del tema...	Pretest	111,546	3	0,000 0,000	No se puede calcular t		
		Postest	111,546					
D7 Conoc. de propias estrategias de relación...	Pretest	103,426	3	16,202 16,202	0,000	2	1,000	
	Postest	103,426						
D8 Metaconocimiento estratégico de lectura	Pretest	110,793	3	4,911 12,995	1,732	2	0,225	
	Postest	102,285						
D9 Reconocimiento de cap. de generalización...	Pretest	87,440	3	12,510 12,510	-1,000	2	0,423	
	Postest	101,886						
Dimensiones 8 y 9	Pretest	105,194	3	7,539 11,516	0,378	2	0,742	
	Postest	102,681						

Grupos	Dimensiones		Media	N	DT	t	gl	Sign.
Grupo F CEIP Sánchez Alonso	Puntuación total	Pretest	99,323	8	13,386 11,707	0,213	7	0,838
		Postest	98,439					
	D1 Conocimiento de las propias limitaciones...	Pretest	96,095	8	19,008 14,433	0,814	7	0,442
		Postest	91,032					
	D2 Conocimiento de las propias expectativas...	Pretest	94,993	8	14,376 16,243	-0,333	7	0,749
		Postest	96,955					
	D3 Conocimiento de la propia actitud...	Pretest	104,461	8	19,916 14,790	0,778	7	0,462
		Postest	99,269					
	D4 Conocimiento de los propios intereses...	Pretest	102,765	8	12,064 16,910	0,104	7	0,920
		Postest	102,059					
	D5 Percepción del apoyo del contexto...	Pretest	99,860	8	10,827 18,017	0,107	7	0,918
		Postest	99,014					
	D6 Percepción de la relevancia del tema...	Pretest	95,970	8	18,617 11,589	-1,357	7	0,217
		Postest	105,705					
D7 Conoc. de propias estrategias de relación...	Pretest	101,400	8	18,244 17,185	1,323	7	0,227	
	Postest	93,299						
D8 Metaconocimiento estratégico de lectura	Pretest	98,386	8	11,538 13,642	-2,049	7	0,080	
	Postest	107,957						
D9 Reconocimiento de cap. de generalización...	Pretest	97,372	8	18,083 13,887	0,683	7	0,516	
	Postest	91,954						
Dimensiones 8 y 9	Pretest	97,655	8	9,010 11,705	-1,313	7	0,231	
	Postest	104,252						

Tabla 51. Comparaciones pretest-postest para la escala APSL por grupos de aplicadores
(Continuación) Estadísticos de muestras relacionadas

Grupos	Dimensiones		Media	N	DT	t	gl	Sign.
Grupo G IES Leonor de Guzmán	Puntuación total	Pretest	98,871	9	17,498 15,780	0,190	8	0,854
		Postest	98,400					
	D1 Conocimiento de las propias limitaciones...	Pretest	97,783	9	17,860 13,072	0,000	8	1,000
		Postest	97,783					
	D2 Conocimiento de las propias expectativas...	Pretest	100,442	9	17,150 20,046	0,000	8	1,000
		Postest	100,442					
	D3 Conocimiento de la propia actitud...	Pretest	97,786	9	16,781 16,723	0,478	8	0,645
		Postest	96,467					
	D4 Conocimiento de los propios intereses...	Pretest	92,734	9	19,088 18,737	-0,336	8	0,746
		Postest	93,988					
	D5 Percepción del apoyo del contexto...	Pretest	105,601	9	16,168 16,596	2,204	8	0,059
		Postest	92,803					
	D6 Percepción de la relevancia del tema...	Pretest	104,623	9	11,316 11,316	0,000	8	1,000
		Postest	104,623					
D7 Conoc. de propias estrategias de relación...	Pretest	107,026	9	17,708 8,101	-1,315	8	0,225	
	Postest	114,227						
D8 Metaconocimiento estratégico de lectura	Pretest	94,723	9	17,596 10,515	-2,268	8	0,053	
	Postest	103,230						
D9 Reconocimiento de cap. de generalización...	Pretest	104,294	9	15,742 11,420	-0,426	8	0,681	
	Postest	106,702						
Dimensiones 8 y 9	Pretest	96,818	9	18,637 11,920	-2,294	8	0,051	
	Postest	105,194						

Grupos	Dimensiones		Media	N	DT	t	gl	Sign.
Grupo H IES Leonor de Guzmán	Puntuación total	Pretest	102,683	4	12,513 16,717	0,253	3	0,817
		Postest	100,207					
	D1 Conocimiento de las propias limitaciones...	Pretest	94,408	4	20,251 19,487	0,000	3	1,000
		Postest	94,408					
	D2 Conocimiento de las propias expectativas...	Pretest	104,801	4	11,096 16,175	0,225	3	0,836
		Postest	102,839					
	D3 Conocimiento de la propia actitud...	Pretest	88,886	4	10,276 7,659	-1,549	3	0,219
		Postest	100,752					
	D4 Conocimiento de los propios intereses...	Pretest	105,586	4	17,542 11,632	0,313	3	0,775
		Postest	101,354					
	D5 Percepción del apoyo del contexto...	Pretest	115,952	4	12,825 19,946	1,093	3	0,354
		Postest	104,095					
	D6 Percepción de la relevancia del tema...	Pretest	99,864	4	7,788 12,718	0,397	3	0,718
		Postest	95,970					
D7 Conoc. de propias estrategias de relación...	Pretest	111,527	4	16,202 13,229	1,732	3	0,182	
	Postest	87,223						
D8 Metaconocimiento estratégico de lectura	Pretest	101,576	4	8,145 18,378	0,264	3	0,809	
	Postest	99,449						
D9 Reconocimiento de cap. de generalización...	Pretest	94,663	4	17,693 0,000	-2,449	3	0,092	
	Postest	116,332						
Dimensiones 8 y 9	Pretest	99,540	4	11,308 16,286	-1,192	3	0,319	
	Postest	105,194						

Tabla 51. Comparaciones pretest-postest para la escala APSL por grupos de aplicadores
(Continuación) Estadísticos de muestras relacionadas

Grupos	Dimensiones		Media	N	DT	t	gl	Sign.
Grupo I IES Mariana Pineda	Puntuación total	Pretest	90,730	5	14,508 17,735	0,162	4	0,879
		Postest	89,881					
	D1 Conocimiento de las propias limitaciones...	Pretest	92,382	5	19,330 19,448	-0,206	4	0,847
		Postest	93,733					
	D2 Conocimiento de las propias expectativas...	Pretest	101,662	5	13,129 20,004	1,725	4	0,160
		Postest	89,108					
	D3 Conocimiento de la propia actitud...	Pretest	89,479	5	16,570 14,289	0,000	4	1,000
		Postest	89,479					
	D4 Conocimiento de los propios intereses...	Pretest	88,094	5	6,433 7,356	1,177	4	0,305
		Postest	84,709					
	D5 Percepción del apoyo del contexto...	Pretest	93,254	5	19,520 14,844	-1,826	4	0,142
		Postest	100,030					
	D6 Percepción de la relevancia del tema...	Pretest	99,085	5	13,032 13,032	0,000	4	1,000
		Postest	99,085					
D7 Conoc. de propias estrategias de relación...	Pretest	106,666	5	21,125 16,202	0,408	4	0,704	
	Postest	103,426						
D8 Metaconocimiento estratégico de lectura	Pretest	90,942	5	17,015 18,639	-0,431	4	0,688	
	Postest	94,345						
D9 Reconocimiento de cap. de generalización...	Pretest	98,997	5	18,130 11,868	-1,000	4	0,374	
	Postest	103,331						
Dimensiones 8 y 9	Pretest	91,624	5	19,514 17,999	-0,535	4	0,621	
	Postest	96,148						

Grupos	Dimensiones		Media	N	DT	t	gl	Sign.
Grupo J IES Torreblanca	Puntuación total	Pretest	87,476	4	4,967 21,495	-0,686	3	0,542
		Postest	94,903					
	D1 Conocimiento de las propias limitaciones...	Pretest	91,032	4	13,501 8,494	-0,167	3	0,878
		Postest	92,720					
	D2 Conocimiento de las propias expectativas...	Pretest	81,262	4	16,949 17,398	-1,000	3	0,391
		Postest	91,070					
	D3 Conocimiento de la propia actitud...	Pretest	77,019	4	10,276 13,156	-2,201	3	0,115
		Postest	102,236					
	D4 Conocimiento de los propios intereses...	Pretest	104,175	4	5,402 14,199	1,578	3	0,213
		Postest	94,301					
	D5 Percepción del apoyo del contexto...	Pretest	100,707	4	17,051 27,936	0,000	3	1,000
		Postest	100,707					
	D6 Percepción de la relevancia del tema...	Pretest	95,970	4	22,028 26,602	0,676	3	0,547
		Postest	84,287					
D7 Conoc. de propias estrategias de relación...	Pretest	95,324	4	16,202 22,914	-1,732	3	0,182	
	Postest	103,426						
D8 Metaconocimiento estratégico de lectura	Pretest	88,815	4	12,761 14,529	-1,095	3	0,353	
	Postest	97,322						
D9 Reconocimiento de cap. de generalización...	Pretest	94,663	4	17,693 21,669	-1,732	3	0,182	
	Postest	105,498						
Dimensiones 8 y 9	Pretest	88,231	4	7,218 16,716	-1,567	3	0,215	
	Postest	99,540						

Tabla 51. Comparaciones pretest-postest para la escala APSL por grupos de aplicadores
(Continuación) Estadísticos de muestras relacionadas

Grupos	Dimensiones		Media	N	DT	t	gl	Sign.
Grupo K IES Torreblanca	Puntuación total	Pretest	106,573	4	11,375	0,293	3	0,789
		Postest	106,219		11,968			
	D1 Conocimiento de las propias limitaciones...	Pretest	99,471	4	18,589	-1,321	3	0,278
		Postest	107,909		20,251			
	D2 Conocimiento de las propias expectativas...	Pretest	100,878	4	18,677	0,522	3	0,638
		Postest	94,993		13,400			
	D3 Conocimiento de la propia actitud...	Pretest	103,719	4	14,533	2,611	3	0,080
		Postest	96,302		8,899			
	D4 Conocimiento de los propios intereses...	Pretest	99,943	4	14,928	No se puede calcular t		
		Postest	105,586		14,928			
	D5 Percepción del apoyo del contexto...	Pretest	117,646	4	7,823	0,926	3	0,423
		Postest	110,871		9,582			
	D6 Percepción de la relevancia del tema...	Pretest	107,652	4	7,788	-1,000	3	0,391
		Postest	111,546		0,000			
D7 Conoc. de propias estrategias de relación...	Pretest	107,476	4	15,512	0,522	3	0,638	
	Postest	103,426		13,229				
D8 Metaconocimiento estratégico de lectura	Pretest	107,957	4	12,031	0,522	3	0,638	
	Postest	105,830		8,145				
D9 Reconocimiento de cap. de generalización...	Pretest	89,246	4	20,746	-1,567	3	0,215	
	Postest	105,498		12,510				
Dimensiones 8 y 9	Pretest	103,310	4	12,875	-0,775	3	0,495	
	Postest	107,079		7,218				

La exploración de los datos nos ofrece información sobre algunos aspectos de interés con relación a nuestro estudio. Por una parte, hemos de constatar que no se han encontrado diferencias significativas a favor del GE en ninguno de los subgrupos estudiados, ni para la puntuación global ni para las dimensiones consideradas. Como podemos observar, las diferencias son notables en ocasiones, lo que no se muestra suficiente para obtener valores de significación estadística, debido probablemente al tamaño exiguo de la muestra, unido a la importante variabilidad de las puntuaciones obtenidas. Por lo que respecta a las diferencias encontradas entre los grupos de aplicación, podemos destacar que en primaria, los subgrupos C, D y, en menor medida, también el subgrupo E, obtienen ganancias en la puntuación global, mientras que los tres grupos restantes muestran pérdidas más o menos pronunciadas. En Secundaria, es el subgrupo J quien se destaca por los resultados positivos en cuanto a la puntuación global de la escala, mientras que los demás subgrupos mantienen sus puntuaciones sin cambios apreciables. Por lo que respecta a las dimensiones de la escala, lo más destacable es que precisamente son las dimensiones D8, D9 y D8 y 9, es decir, aquellas directamente relacionadas con los procesos metacognitivos en la lectura, junto con la D1 (conocimiento de las propias limitaciones cognitivas actuales, potencialidades y posibilidades de aprendizaje, dificultades y vías de solución), las que obtienen un mayor número de ganancias entre los distintos subgrupos. En efecto, la D1 obtiene 7 ganancias; la D8 obtiene 6; 8 son los subgrupos que obtienen ganancias postest en la D9 y un total de 9 los que muestran ganancias en la dimensión conjunta D8 y 9. Posteriormente, cuando abordemos el análisis y la discusión de los resultados obtenidos, nos adentraremos en las posibles causas que podrían explicar el hecho de haber obtenido un mayor número de ganancias postratamiento precisamente en estas dimensiones directamente relacionadas con el metaconocimiento estratégico sobre la lectura.

(b) Criterio del profesorado + puntuación inicial en la escala APSL

Analizaremos a continuación los resultados en función de este segundo criterio. Como apuntábamos anteriormente, unido a la designación previa del profesorado, hemos considerado la obtención de una puntuación baja (menor o igual a 84) en la escala, como otro factor relevante a la hora de establecer comparaciones pretest-postest y valorar las posibles mejoras obtenidas por los grupos. Los resultados obtenidos por los grupos experimental y control, así como el valor de los contrastes, los exponemos en la siguiente tabla.

Tabla 52. Comparación pretest-postest para la escala APSL en función del criterio de designación del profesorado más la puntuación inicial en la prueba (puntuación ≤ 84)

Prueba t para muestras relacionadas

Gr.	Dimensiones	N	Pretest	DT	Postest	DT	T	gl	sign.
GE GC	Puntuación global	11 12	78,024 75,334	3,991 9,294	86,383 79,342	9,554 13,189	-3,609 -1,082	10 11	0,005** 0,303
GE GC	D1 Conocimiento de las propias limitaciones...	11 12	78,758 82,594	8,441 9,159	91,032 87,657	13,501 14,274	-3,194 -1,431	10 11	0,010** 0,180
GE GC	D2 Conocimiento de las propias expectativas...	11 12	89,108 77,993	11,096 13,979	84,829 84,531	13,298 15,132	0,971 -1,520	10 11	0,355 0,157
GE GC	D3 Conocimiento de la propia actitud...	11 12	76,210 87,402	12,800 16,420	86,998 89,380	14,244 17,057	-2,319 -0,432	10 11	0,043* 0,674
GE GC	D4 Conocimiento de los propios intereses...	11 12	88,402 80,195	8,506 7,217	91,480 81,135	14,049 9,573	-0,629 -0,394	10 11	0,543 0,701
GE GC	D5 Percepción del apoyo del contexto...	11 12	90,544 79,816	17,141 11,720	87,464 85,462	15,850 17,364	0,518 -0,989	10 11	0,616 0,344
GE GC	D6 Percepción de la relevancia del tema...	11 12	86,057 94,672	20,036 18,139	94,554 88,181	14,702 26,149	-1,322 1,239	10 11	0,216 0,241
GE GC	D7 Conoc. de propias estrategias de relación...	11 12	93,115 85,873	18,148 12,848	104,898 88,573	16,923 16,141	-1,620 -0,561	10 11	0,136 0,586
GE GC	D8 Metaconocimiento estratégico de lectura	11 12	83,208 82,434	14,465 14,957	91,715 83,143	15,895 12,804	-1,546 -0,146	10 11	0,153 0,886
GE GC	D9 Reconocimiento de cap. de generalización...	11 12	94,663 96,469	16,785 17,183	104,513 98,275	11,316 12,510	-1,614 -0,432	10 11	0,138 0,674
GE GC	Dimensiones 8 y 9	11 12	83,263 83,205	14,086 13,041	94,228 84,462	15,214 13,303	-2,451 -0,257	10 11	0,034* 0,802

El análisis de las comparaciones efectuadas muestra ganancias muy significativas en el GE para la puntuación global de la escala; así mismo también se observan diferencias con significatividad estadística al nivel 0,01 para la dimensión D1, y al nivel de significación 0,05 para las dimensiones D3 y D8 y 9. Podemos observar que al resultar significativo el contraste para la dimensión D8 y 9, las comparaciones para las dimensiones que conforman ésta, es decir, D8 y D9, obtienen unos valores altos en cuanto a la diferencia de medias pretest-postest. El GC1, en cambio, aunque experimenta una subida en sus puntuaciones postest para la puntuación global y para algunas de las dimensiones, no alcanza la significatividad en ninguno de los contrastes efectuados.

Hemos querido realizar un contraste de las diferencias de medias obtenidas por los grupos experimental y control. A pesar de que los resultados expuestos indican de forma fehaciente una ganancia significativa del GE y no así para el GC1, el contraste de diferencias promedio puede matizar estos resultados, ya que podría suceder que la variación real entre ambos grupos no fuera tan grande, pero sí lo suficiente como para otorgar a uno significatividad y a otro no. El contraste realizado muestra la ausencia de diferencias estadísticamente significativas entre las diferencias de medias de ambos grupos para la puntuación global, a pesar de que el valor de la diferencia ha sido de 4,350 puntos. El valor t observado fue de 0,975 y la significación asociada de 0,341 para 21 gl. Por otra parte, los contrastes para las dimensiones D6, D7, D8, D9 y D8 y 9, expresan valores de diferencia que van desde los 8 hasta los 15 puntos, acercándose todos ellos a la significación estadística.

El análisis de regresión estadística nos ofrece la posibilidad de evaluar en qué medida los resultados positivos y diferenciales obtenidos por el GE en el postest de la escala, son debidos a las puntuaciones obtenidas inicialmente. Realizados los cálculos, hemos extraído una correlación entre las puntuaciones pretest y postest de 0,63. Este resultado nos ofrece la posibilidad de calcular la bondad de ajuste, es decir, cómo de bien explica la puntuación inicial, la obtenida finalmente; ésta quedaría expresada de la siguiente forma: $R^2 = 0,39$ (R^2 corregida = 0,33), mientras que la prueba F para evaluar la significación del contraste arrojó un valor de 5,985, con una significación asociada de 0,037. De esta forma, como es esperable por otra parte, la puntuación inicial explicaría una parte significativa de la puntuación postratamiento. Es precisamente la constatación de que existe proporción de varianza no explicada por el pretest, lo que induce a pensar en la efectividad del tratamiento aplicado y en la información adicional aportada, que posiblemente haya que extraer del análisis del proceso de aplicación del dispositivo experimental.

Igualmente nos ha interesado evaluar cuál ha sido el comportamiento de los sujetos que habiendo sido designados por sus profesores u orientadores como alumnos con dificultades de aprendizaje de la lectura, obtuvieron una baja puntuación en ECO 1 o ECOS, es decir, igual o inferior a 84 puntos. Esto interesaba especialmente dado que la investigación tiene como eje principal la evaluación de la comprensión de textos, por lo que queríamos conocer cómo habían realizado la escala APSL aquellos que habían mostrado mayores dificultades en la prueba criterio de comprensión. Los resultados en primaria muestran para el GE (N = 9) una media pretest de 95,256 (DT = 10,151) y una postest de 97,928 (DT = 8,382); el valor t observado fue de -0,745 y la significación de 0,478 para 8 gl. En cuanto a las dimensiones de la escala, no hemos encontrado diferencias significativas para ninguna de ellas, aunque sí se hallan diferencias importantes a favor de la puntuación postratamiento en la D3, D4 y D8. En cuanto al GC1 (N = 10), la puntuación media pretest fue de 97,802 (DT = 19,689) y la postest de 97,661 (DT = 15,089), con un valor t de 0,029 y una significación de 0,977 para 9 gl. No se encontraron diferencias estadísticamente significativas para ninguna de las dimensiones de la escala.

En cuanto a secundaria, para el GE (N = 4) la media pretest fue de 95,963 (DT = 16,190) y la postest de 100,561 (DT = 14,210), con un valor t de -1,253 y una significación de 0,299 para 3 gl. En cuanto a las dimensiones, hemos encontrado diferencias significativas al nivel 0,05 para la D3. El GC1 (N = 3) obtuvo una media pretest de 92,427 (DT = 9,276) y una postest de 80,167 (DT = 14,449), con una t de

3,606 y una significación de 0,069 para 2 gl. En general las diferencias halladas para las dimensiones de la escala expresan pérdidas posttest, aunque tan sólo una de ellas, la D4, alcanza la significatividad estadística.

También hemos realizado los cálculos para los sujetos que además de obtener una puntuación baja en las pruebas de comprensión, habían obtenido igualmente bajos resultados (≤ 84) en la propia escala. En primaria, para el GE ($N = 2$), la media pretest fue de 79,696 (DT = 0,000) y la posttest de 96,671 (DT = 4,001), con un valor t de -6,000 y una significación de 0,105 para 1 gl. Se dan grandes diferencias en algunas de las dimensiones a favor de la puntuación posttest, que se acercan a la significatividad estadística, sobre todo para las dimensiones D3 y D4. El contraste para el GC1 no pudo realizarse debido a que tan sólo un sujeto presentaba las condiciones del criterio formulado. El único sujeto válido para el contraste había obtenido una ganancia cercana a los veinte puntos. En lo que respecta al GE de Secundaria ($N = 2$), la media pretest fue de 82,525 (DT = 0,000), mientras que la puntuación media posttest fue de 90,305 (DT = 11,002), con un valor t de -1,000 y una significación de 0,500 para 1 grado de libertad. No había diferencias estadísticamente reseñables para las dimensiones, a pesar de que las puntuaciones postratamiento para estos dos sujetos eran muy superiores a las iniciales. En cuanto al GC1, al igual que ocurriera con la submuestra de primaria, no se pudieron realizar los cálculos estadísticos debido a que sólo uno de los alumnos permaneció en el grupo cuando se aplicó este criterio. Los resultados para este alumno indicaban una pérdida en la puntuación posttest de unos 14 puntos.

(c) Criterio del profesorado + puntuación inicial en la escala APSL + puntuación inicial en Cattell

También en este contraste, como ya ocurriera en función del segundo de los criterios adoptados, las diferencias pretest-posttest para la puntuación global de la escala se muestran significativas estadísticamente a favor del GE, no siendo así, a pesar de las diferencias de medias observadas, para el GC1. Además de la significatividad del contraste para la puntuación global en el GE, también se dan diferencias muy significativas en el contraste para la dimensión D1, y significativas al nivel 0,05 para las dimensiones D9 y D8 y 9. Igualmente, podemos observar que para otras dimensiones de las contempladas en la prueba, los resultados expresan notables diferencias que se acercan a la significatividad estadística. Por el contrario, hay que reseñar que en la dimensión D2, las diferencias son a favor de las puntuaciones pretest, llegando éstas a ser significativas al nivel 0,05. Por lo que respecta al GC1 y a las dimensiones de la escala, tan sólo en la dimensión D2 se hallan diferencias significativas al nivel 0,05, aunque pueden observarse incrementos posttest importantes para varias dimensiones.

El contraste entre las diferencias de medias de los grupos experimental y control, puede proporcionarnos una información complementaria y matizada de los resultados mostrados. En este sentido, los resultados de este contraste para la puntuación global expresan un valor de diferencia de 2,702 puntos favorables paradójicamente al GC1, con un valor t de 0,522 y una significación asociada de 0,611 para 13 gl. Por otra parte, a pesar de que la mayoría de los contrastes para las dimensiones resultan favorables al GE y alguno de ellos está cercano a la significatividad estadística con valores de diferencia de hasta 13 puntos, es uno de los contrastes favorables al GC1, el realizado para la dimensión D2, el que resulta significativo, con un valor de diferencia de más de 20 puntos, un valor t de 4,065 y una significación de 0,001 para los 13 gl.

Tabla 53. Comparación pretest-postest para la escala APSL en función del criterio de designación del profesorado más la puntuación inicial en la prueba (puntuación ≤ 84) más la puntuación inicial en Cattell (puntuación ≤ 84).

Prueba t para muestras relacionadas

Gr.	Dimensiones	N	Pretest	DT	Postest	DT	T	gl	Sign.
GE GC	Puntuación global	8 7	77,574 75,048	4,277 10,191	85,177 85,354	10,651 9,935	-2,817 -2,232	7 6	0,026* 0,067
GEG C	D1 Conocimiento de las propias limitaciones...	8 7	77,531 82,353	9,546 11,505	91,032 87,175	9,546 16,924	-4,000 -0,918	7 6	0,005** 0,394
GE GC	D2 Conocimiento de las propias expectativas...	8 7	91,070 76,779	11,675 9,983	81,262 91,350	11,096 14,119	2,376 -3,357	7 6	0,049(-) 0,015*
GE GC	D3 Conocimiento de la propia actitud...	8 7	75,536 89,309	13,732 19,635	85,919 97,786	16,479 11,361	-1,618 -1,549	7 6	0,150 0,172
GE GC	D4 Conocimiento de los propios intereses...	8 7	90,069 77,777	9,417 7,179	87,248 82,613	11,969 8,530	0,642 -1,686	7 6	0,542 0,143
GE GC	D5 Percepción del apoyo del contexto...	8 7	93,932 81,833	17,742 10,865	93,085 87,640	14,905 16,987	0,107 -0,666	7 6	0,918 0,534
GE GC	D6 Percepción de la relevancia del tema...	8 7	84,287 95,970	19,965 12,718	90,129 95,970	14,280 23,793	-0,814 0,000	7 6	0,442 1,000
GE GC	D7 Conoc. de propias estrategias de relación...	8 7	89,248 87,223	18,244 16,202	105,451 89,537	18,244 14,578	-2,000 -0,354	7 6	0,086 0,736
GE GC	D8 Metaconocimiento estratégico de lectura	8 7	78,180 78,788	13,642 16,184	90,942 88,511	16,396 12,727	-2,121 -1,922	7 6	0,072 0,103
GE GC	D9 Reconocimiento de cap. de generalización...	8 7	91,954 97,759	13,887 14,953	105,498 103,950	11,582 11,582	-3,416 -1,549	7 6	0,011* 0,172
GE GC	Dimensiones 8 y 9	8 7	77,865 80,423	12,047 14,897	93,886 91,193	17,097 11,863	-3,188 -2,085	7 6	0,015* 0,082

Hemos efectuado un análisis de regresión para la puntuación postest de APSL en función de las puntuaciones iniciales contempladas en el criterio formulado, es decir, en APSL y en Cattell. Cuando fueron consideradas conjuntamente las puntuaciones iniciales de ambas variables, la R fue de 0,819 y la R^2 de 0,671 (R^2 corregida, de 0,539). El valor F de la regresión fue de 5,096, con una significación de 0,062. En relación al ajuste del modelo, debemos concluir que éste no predice adecuadamente los resultados postest en la escala, aunque los datos indican proximidad a la significatividad estadística. Los resultados del análisis muestran una correlación del pretest de Cattell con el pretest de APSL de -0,201 (sig.0,316) y de -0,303 con el postest, con una significación asociada de 0,233. En este sentido, cuando se incluyó la variable independiente relativa a la puntuación CI pretest, la bondad de ajuste experimentó un decremento con respecto al contraste realizado exclusivamente con la puntuación pretest de APSL. Cuando la regresión se realizó sólo con la variable pretest de APSL, el valor F fue de 11,147 y la significación de 0,016; la R^2 fue de 0,650 (R^2 corregida 0,592). Cuando se introdujo en el análisis sólo la puntuación inicial obtenida en Cattell, la R^2 fue de 0,091 y una vez corregida de -0,059. El valor F marcó 0,608 y la significación de este valor 0,465.

Relación entre el ajuste personal-social y la comprensión de textos

Uno de los análisis previstos para valorar los cambios con respecto a los aspectos evaluados en APSL consiste en examinar la correlación existente entre la escala y las pruebas de comprensión, determinando si existe en la fase postest del estudio un mayor ajuste entre las puntuaciones de ambas pruebas; esto ofrecería un indicador valioso de mayor ajuste personal-social con relación a la lectura, controlando la incidencia que las variables expuestas anteriormente podrían tener y evitando un posible sesgo que se produciría en una apreciación simple de los datos.

Hemos procedido a evaluar la evolución de esta correlación entre las puntuaciones de ambas pruebas en las fases pretest y postest, con el fin de determinar si existe un mayor ajuste entre las puntuaciones para el GE con posterioridad al tratamiento. Hemos extraído los datos de las correlaciones para las distintas submuestras formadas en función de los tres criterios formulados. Los resultados para la puntuación global del GE están expuestos en la siguiente tabla.

Tabla 54. Correlaciones entre las medidas de ajuste personal-social y nivel de comprensión en el GE para las fases pretest y postest del estudio

	Criterios	Pretest ECO 1	Postest ECO 1	Pretest ECOS	Postest ECOS
Pretest/Postest APSL	(a) Designación del profesorado	N = 30 r. 0,078 sig. 0,681	N = 28 r. 0,150 sig. 0,446	N = 27 r. 0,061 sig. 0,763	N = 24 r. 0,084 sig. 0,695
	(b) Designación + APSL <= 84	N = 5 r. -0,728 sig. 0,163	N = 5 r. 0,330 sig. 0,587	N = 7 r. -0,530 sig. 0,221	N = 5 r. -0,474 sig. 0,420
	(c) Designación + APSL <= 84 + CI <= 84	N = 3 r. -0,839 sig. 0,367	N = 3 r. -0,070 sig. 0,955	N = 6 r. -0,467 sig. 0,351	N = 4 r. -0,593 sig. 0,407

Podemos observar que para el primero de los criterios contemplados, las correlaciones establecidas experimentan un ligero incremento en la comparación que efectuamos con las puntuaciones postest. Este incremento se produce para las dos pruebas de comprensión elaboradas. En cuanto al segundo de los criterios, para la prueba ECO 1 se produce la mayor evolución en la magnitud de la correlación, que pasa de un valor negativo entre las puntuaciones iniciales de ajuste y comprensión, a un valor de r. 0,330. Este resultado es coherente con los anteriormente expuestos durante el análisis de las puntuaciones de ambas pruebas, en las que fueron halladas diferencias significativas para el GE cuando los sujetos cumplían los dos primeros criterios. En cambio, para el grupo de alumnos de secundaria que cumplían estos dos primeros criterios, la evolución de la correlación, aunque es positiva, no deja entrever un cambio sustancial en la relación que mantienen ambas pruebas; la correlación en el postest sigue siendo negativa entre ajuste y comprensión. Por lo que respecta al tercero de los criterios considerados, lo primero que hay que reseñar es lo exiguo de las muestras, lo que hace difícil una interpretación adecuada de los resultados. Éstos muestran un incremento notable de la correlación para las puntuaciones de ajuste y ECO 1, expresada en la

disminución acentuada de la correlación negativa inicial; mientras que para la prueba de secundaria, se vuelve a manifestar una correlación negativa que ni siquiera experimenta descenso, sino que incrementa de forma ligera su valor. Los resultados de estos contrastes para las pruebas de ajuste y comprensión en la etapa de secundaria alertan contra una falta notable de relación entre ambas puntuaciones, lo que nos conduce a la necesidad de profundizar en el análisis cualitativo de los resultados obtenidos en ambas pruebas, así como de sondear otras posibilidades relativas a variables no controladas que hubieran podido incidir en estos resultados.

Hemos realizado los cálculos para el GC1 en función de los mismos criterios contemplados anteriormente. Los resultados se muestran en la tabla siguiente.

Tabla 55. Correlaciones entre las medidas de ajuste personal-social y nivel de comprensión en el GC1 para las fases pretest y postest del estudio

	Criterios	Pretest ECO 1	Postest ECO 1	Pretest ECOS	Postest ECOS
Pretest/Postest APSL	(a) Designación del profesorado	N = 23 r. 0,134 sig. 0,541	N = 20 r. -0,079 sig. 0,742	N = 37 r. 0,318 sig. 0,055	N = 33 r. 0,181 sig. 0,314
	(b) Designación + APSL <= 84	N = 4 r. 0,156 sig. 0,844	N = 3 r. -0,189 sig. 0,879	N = 9 r. -0,744 sig. 0,021*	N = 9 r. 0,411 sig. 0,272
	(c) Designación + APSL <= 84 + CI <= 84	N = 2 r. -1,000 sig. 0,000**	N = 1 ----- -----	N = 6 r. -0,788 sig. 0,062	N = 6 r. 0,617 sig. 0,192

Los resultados para el GC1 dejan entrever una evolución negativa para la correlación entre las puntuaciones de ajuste y comprensión con relación especialmente a la prueba ECO 1, en la que las correlaciones postest se tornan negativas, tanto en función del primero de los criterios contemplados, como para el segundo de estos; para el tercer criterio, no se ha podido realizar el cálculo debido a que tan sólo uno de los sujetos presentaba las condiciones exigidas. En cambio, para la etapa de secundaria, la evolución de las correlaciones difiere en función de los criterios considerados; así, para el primero, la evolución es igualmente negativa, mientras que para los sujetos que cumplían los dos primeros criterios, así como para aquellos que cumplían los tres, la correlación que se establece entre las puntuaciones postest de las pruebas experimenta un crecimiento notable en el postest con relación a la que se estableció para las puntuaciones iniciales.

Hemos procedido a revisar estos contrastes en función de los sujetos que presentaban dificultades de aprendizaje y recibían AI, o presentaban discapacidad, deprivación social o participaban de programas de DC. La composición de este grupo reducido, como hemos venido comprobando a lo largo de los distintos análisis efectuados, quedaba prácticamente igual para el GE, lo que ha ocasionado que los resultados de las correlaciones sean casi idénticas, destacando que para el primero de los criterios quedan igualadas las magnitudes de las correlaciones para la prueba ECO 1. Para el resto de contrastes, los resultados son los mismos, ya que el número de sujetos no varió cuando se aplicó esta condición. En cambio, para el GC1 se producen cambios notables cuando

se reducen sus efectivos a tenor de las condiciones de mayor grado de problemas de aprendizaje. De esta forma, para el primero de los criterios se halló una evolución positiva de las correlaciones para las prueba ECO 1, que fue desde $r. -0,012$ a $r. 0,323$; en cambio, para la prueba ECOS se acentuó una evolución negativa de la correlación: $r. 0,372$ a $r. 0,008$; para el segundo criterio, la evolución positiva con relación a la prueba ECOS decrece en su intensidad, pasando de $r. -0,106$ a $r. 0,126$; y así mismo, para la misma prueba y en función del tercero de los criterios, la evolución de la correlación pasa a ser muy moderada, pasando de una $r. -0,500$ entre las puntuaciones pretest a $r. -0,337$ entre las postest. Por tanto, podríamos concluir que cuando se igualan las condiciones en cuanto al grado de dificultad de los sujetos que han conformado los grupos, la evolución de las correlaciones con relación a la prueba ECOS, aún siendo positiva, se torna mucho más moderada.

Hemos recogido en la siguiente tabla una síntesis de la evolución de las correlaciones; para ello hemos procedido sencillamente a restar los valores absolutos de las correlaciones pretest y postest para cada una de las pruebas y en función de los tres criterios. A pesar de lo expresado anteriormente con relación a los efectivos del GC1, los datos corresponden nuevamente a los grupos completos.

Tabla 56. Evolución de las correlaciones pretest/postest entre APSL y las pruebas de comprensión para los grupos experimental y control en función de los tres criterios contemplados

	Criterios	Grupo experimental		Grupo control	
		Evolución positiva	Evolución negativa	Evolución positiva	Evolución negativa
Pretest/Postest APSL/ECO1	(a)	+ 0,072			- 0,213
	(a) + (b)	+ 1,058			- 0,345
	(a) + (b) + (c)	+ 0,769			No es posible
Pretest/Postest APSL/ECOS	(a)	+ 0,023			- 0,137
	(a) + (b)	+ 0,056		+ 1,155	
	(a) + (b) + (c)		- 0,126	+ 1,405	

Podemos observar que en lo referente a la prueba de comprensión ECO 1, la evolución de las correlaciones ha sido positiva en todos los criterios para el GE, mostrándose particularmente amplia para el criterio que hacía referencia a la designación del profesorado junto con una baja puntuación en la escala. En cambio, para el GC1, la evolución ha tenido un matiz claramente negativo. No sucede lo mismo con relación a la etapa de secundaria. Como apuntábamos anteriormente, la evolución de las correlaciones APSL/ECOS para el GE no ha sido tan positiva como cabía esperar a tenor de los resultados mostrados en ambas pruebas por este grupo; en cambio, el GC1 ha mostrado una evolución positiva de gran intensidad, pasando de correlaciones negativas en el pretest a otras ampliamente positivas en el postest.

Las DDP en la escala APSL

El análisis con el instrumento exploratorio de las DDP que hemos efectuado con las pruebas de comprensión, si bien aplicado sólo en función de uno de los criterios formulados, nos ha permitido conocer cuántos sujetos obtuvieron ganancias y cuántos pérdidas en cada uno de los subgrupos —experimental o control— comparados.

Igualmente nos permitió acceder a la magnitud de las ganancias o pérdidas manifestadas. Nos interesa aplicar este instrumento a la escala de ajuste personal-social por estos mismos motivos. Al igual que con las pruebas de comprensión, hemos optado por realizar la aplicación de las DDP al contraste efectuado con el grupo de sujetos que precisaba atención educativa especial, es decir, sujetos con dificultades de aprendizaje, privación social o discapacidad que recibían AI o participaban de programas de DC. En este caso, el GE estaba formado por 52 sujetos. Hemos procedido a calcular en primer lugar el número de sujetos que obtuvieron ganancias o pérdidas en el postest. Los resultados muestran que 24 alumnos obtuvieron ganancias postest, representando el 46,15% de la muestra y 28 los que obtuvieron pérdidas, de los que 4 mostraron una ganancia nula; en total, representaban el 53,85%. Hemos otorgado a continuación a cada sujeto una puntuación que se correspondía con la magnitud de sus ganancias o pérdidas y hemos vuelto a convertir en porcentaje la suma de los puntos ganadores y perdedores obtenidos, calculando finalmente la puntuación porcentual a favor y en contra de hipótesis. Los resultados se muestran a continuación.

	Ganancias	Pérdidas	A favor de hipótesis %	En contra de hipótesis %
Sujetos	24	28	46,15	53,85
Puntos	229,16	237,41	49,12	50,88

Las puntuaciones obtenidas con estos primeros cálculos ilustran lo que ya habíamos expuesto mediante los análisis psicométricos efectuados anteriormente; las diferencias entre ganancias y pérdidas en el GE no son indicativas de mejora global del grupo en la fase postratamiento. Esto es aplicable, tanto en el número de sujetos, como en la comparación una vez que se ha asignado a cada uno la magnitud de la diferencia pretest-postest conseguida. El contraste bajo este criterio de designación previa del profesorado indica la necesidad de profundizar en el análisis de qué sujetos son los que han obtenido mejoras y por qué. A descubrir esto nos ha ayudado la ejecución de algunos de los contrastes que ya mostramos con posterioridad al realizado con el grupo completo bajo el primero de los criterios. De estos se deducen algunas claves que pueden resultarnos de gran interés en las conclusiones sobre la aplicación del dispositivo EDPL, aunque éstas serán abordadas más adelante, una vez que hayamos introducido en el análisis la evaluación del proceso y de los resultados de la aplicación del tratamiento.

Ahora debemos continuar con la realización del contraste que nos dará la probabilidad con la que pueden ser obtenidos los resultados mostrados. Con la prueba t de Student para muestras independientes obtenemos el valor estadístico y la significación de la comparación entre las puntuaciones de aquellos que obtuvieron pérdidas y las de aquellos que mostraron ganancias. En cuanto a los requisitos para la aplicación de esta prueba, pudimos comprobar en los análisis efectuados, que las puntuaciones obtenidas para la aplicación piloto de la prueba APSL no presentaban una distribución normal, por lo que no podríamos garantizar que la variable *ajuste personal-social* esté distribuida normalmente entre la población; tampoco en un nuevo contraste, realizado éste para el grupo de sujetos con dificultades de aprendizaje con el que se están realizando estas comparaciones, obtuvimos mejores resultados en la prueba. Hemos optado por la aplicación de la prueba t en virtud de que los resultados del contraste de normalidad para la aplicación piloto no se distanciaban mucho de la significación estadística y considerando que la potencia de la prueba puede ofrecer suficientes garantías. Aún así,

creemos conveniente tomar con mayor cautela los resultados del contraste que exponemos en la siguiente tabla.

*Tabla 57. Contraste en función de las DDP para el GE
Prueba t para muestras independientes*

	N	Diferencia media (pre — post)	DT	Prueba t	gl.	Sign.
Grupo	24 (G)	9,548	7,649			
experimental	28 (P)	8,478	7,269	-0,516	50	0,608

El análisis nos muestra que la diferencia promedio pretest-postest es favorable al grupo de ganancias, a pesar de lo cual se mostraba en los resultados anteriores que el porcentaje de sujetos y de puntos era superior en las pérdidas; esta paradoja tiene relación con la asignación de un punto a los sujetos que obtuvieron ganancias nulas, pasando a engrosar la casilla de pérdidas. Por lo demás, las diferencias son reducidas y no expresan significación alguna.

Hemos realizado los cálculos para el GC1, también reducido. En este caso, los sujetos que se mantenían tras la aplicación de este criterio, adoptado con la pretensión de *igualar* ambos grupos, fue de 30. De ellos, 18 integraron la columna de pérdidas, de los cuales 2 habían mostrado una diferencia nula entre pretest y postest, representando en total el 60% de la submuestra. Por otra parte, 12 alumnos obtuvieron ganancias, lo que representaba el 40% del total. Los resultados del GC1 se muestran a continuación.

	Ganancias	Pérdidas	A favor de hipótesis %	En contra de hipótesis %
Sujetos	12	18	40	60
Puntos	106,09	149,12	41,57	58,43

Observamos a través de estos análisis que los porcentajes de sujetos y de puntos que expresan la DDP, muestran con mayor claridad las diferencias entre las puntuaciones pretest-postest entre los grupos experimental y control. Mientras que para el GE las puntuaciones medias analizadas en los contrastes psicométricos eran prácticamente idénticas, lo que también se reflejó en el contraste realizado con las DDP, las diferencias en el GC1 expresan pérdidas, situándose en 20 puntos porcentuales en cuanto a los sujetos y en unos 16 cuando se asignan a éstos las magnitudes de ganancias o pérdidas.

*Tabla 58. Contraste en función de las DDP para el GC1
Prueba t para muestras independientes*

	N	Diferencia media (pre — post)	DT	Prueba t	gl.	Sign.
Grupo	12 (G)	8,841	6,959			
control	18 (P)	8,284	7,684	-0,202	28	0,842

Por su parte, el análisis del contraste efectuado con la prueba t nos muestra, al igual que ya ocurriera con el GE, que la magnitud de las diferencias promedio entre pretest y postest para el grupo que obtuvo ganancias supera a la que fue obtenida por el grupo de

pérdidas, a pesar de que el número de sujetos fuera sensiblemente menor entre los que obtuvieron mejoras posttest. En cualquier caso, y de igual forma que el grupo de tratamiento, la probabilidad asociada a las diferencias observadas carece de significatividad.

Análisis de la fiabilidad de las mejoras en el grupo experimental

Desde la doble perspectiva de las propiedades psicométricas de los instrumentos de evaluación, utilizados para la comprobación de las mejoras, por una parte, así como para el análisis de la validez que pudieran tener las propias mejoras observadas, por otra; se ha descrito que más allá de evaluar la fiabilidad test-retest de una prueba, lo que se muestra realmente necesario es una evaluación de la fiabilidad del cambio observado (Sternberg y Grigorenko, 2003). Se han propuesto diversos métodos para valorar esta propiedad; por ejemplo Carver (1974), citado en Sternberg y Grigorenko (2003), propuso el empleo de dos pretest diferentes seguidos de una fase de intervención muy normalizada y dos posttest a su vez también diferentes. También Embretson (1987) estableció que la utilización de tres posttest diferentes para un mismo pretest, permitiría evaluar repetidamente la fiabilidad de las mejoras obtenidas. En este sentido, nosotros hemos hablado de que la utilización conjunta de la prueba de comprensión de textos y de la escala de ajuste personal-social podría ofrecer ciertos datos que informarían sobre la posible convergencia de los resultados. Si bien estas pruebas no son en absoluto equivalentes, es decir, no tratan de medir los mismos procesos, sí contienen en cambio aspectos y dimensiones similares en su estructura, al tiempo que ambas evalúan procesos implicados en la lectura contemplados en nuestra teoría de base.

Hemos tratado, por tanto, de aplicar, aún de forma limitada, algunos procesos de valoración estadística a los cambios producidos en ambas pruebas, con el fin de evaluar la fiabilidad de las mejoras obtenidas. En este sentido, haremos especial énfasis en la correlación de las dimensiones D8, D9 y D8 y 9 de la escala APSL con las pruebas de comprensión, ya que, como hemos expuesto con anterioridad, el contenido de estas dimensiones hace referencia al conocimiento metacognitivo sobre la comprensión de textos, algo que se encuentra recogido igualmente en las pruebas de comprensión de primaria y secundaria.

En primer lugar, para valorar la posible convergencia de los resultados obtenidos en ambas pruebas, hemos recurrido a correlacionar las puntuaciones obtenidas en las pruebas de comprensión ECO 1 y ECOS con las obtenidas en la puntuación global de la escala —dato éste que ya obtuvimos anteriormente con motivo del análisis de la evolución de las correlaciones entre la escala de ajuste y las pruebas de comprensión—, así como en las dimensiones contempladas, con especial atención a las dimensiones D8, D9 y D8 y 9. Los resultados arrojan en primaria una correlación de 0,150 para las puntuaciones globales de ECO 1 y APSL, con una significación de 0,446 siendo 28 el número de sujetos considerados. Tan sólo para la D2 la correlación es significativa ($r = 0,412$; sig. 0,030). En cambio, para las dimensiones de metacognición, las correlaciones resultan negativas (D8 $r = -0,162$; sig. 0,411; D9 $r = -0,081$; sig. 0,683; y D8 y 9 $r = -0,174$; sig. 0,375). Por otra parte, hemos correlacionado las puntuaciones que expresaban las diferencias pretest-posttest en la escala APSL con la puntuación total de ECO 1. Este contraste ($N = 27$) lo hemos efectuado para las puntuaciones totales y para las dimensiones de metacognición. En este caso, los resultados son significativos para la dimensión D8 ($r = 0,446$; sig. 0,020) y para la D8 y 9 ($r = 0,409$; sig. 0,034). Para la

dimensión D9 la correlación es positiva, aunque no relevante, mientras que para la puntuación global la correlación es de 0,196, con una significación asociada de 0,328.

Hemos efectuado también el contraste entre las puntuaciones obtenidas en las dimensiones de metacognición y las obtenidas en la dimensión *Metaconocimiento estratégico* de la prueba ECO 1. Los resultados expresan correlaciones débiles y negativas para la D8, la D9, así como para la puntuación global, mientras que para la dimensión D8 y 9 la correlación es positiva, aunque igualmente débil ($r=0,063$; sig. 0,752). En cuanto a las correlaciones entre las puntuaciones diferencia para ambas pruebas, los resultados muestran que para la puntuación global de ECO 1 las correlaciones son débiles y negativas con relación a las dimensiones D8, D9, D8 y 9, así como para la puntuación global de la escala ($r=-0,035$; sig. 0,863). El contraste con las diferencias de medias para la dimensión *Metaconocimiento estratégico* en ECO 1, arroja resultados de correlación negativa con relación a la dimensión D8 ($r=-0,068$; sig. 0,737), mientras que para la D9 ($r=0,126$; sig. 0,532), D8 y 9 ($r=0,003$; sig. 0,987), y la puntuación global ($r=0,104$; sig. 0,606), la correlación es positiva, aunque en ningún caso significativa.

Si bien los anteriores contrastes parecen apuntar de forma más o menos clara a la inexistencia o escasez de convergencia entre las puntuaciones y entre las mejoras obtenidas en ambas pruebas, un análisis más exploratorio e intuitivo nos ofrecerá datos más comprensibles. En este sentido, el simple cálculo del número de alumnos que han obtenido mejoras en ambas pruebas nos puede dar una idea de la convergencia del cambio obtenido en el grupo experimental. En el caso de primaria, del total de sujetos que conformaron el GE, tan sólo 5 obtuvieron ganancias en la puntuación global de las dos pruebas. Este número es de 4 cuando lo que se valora es el haber compartido mejoras en la puntuación global de ECO 1 y en la dimensión conjunta D8 y 9 de la escala. Por otra parte, cuando los contrastes los efectuamos seleccionando aquellos sujetos que obtuvieron puntuaciones pretest que se encontraban por debajo de una desviación típica de la media, es decir, iguales o inferiores a 84 puntos, los resultados ($N = 11$) para la correlación entre las diferencias medias para la puntuación global de ECO1 y APSL indican un pequeño pero sustancial cambio, pasando de ser negativas a positivas, aunque débiles ($r=0,061$; sig. 0,885). En cambio, para las dimensiones de metacognición, las correlaciones siguen siendo negativas. Igualmente, para la correlación entre la dimensión *Metaconocimiento estratégico* de ECO 1 y la puntuación global de APSL, la correlación aumenta, pasando de $r=0,104$ a $r=0,217$ con esta submuestra; mientras que nuevamente, para las dimensiones de metacognición de la escala, las correlaciones siguen siendo negativas, incrementándose incluso el valor de algunas de ellas.

Por lo que respecta a secundaria ($N = 25$), la correlación entre las puntuaciones globales de la prueba ECOS y la escala APSL, es positiva aunque débil ($r=0,084$; sig. 0,695). Para las dimensiones de la escala no encontramos correlaciones significativas, y en lo referente a las dimensiones de metacognición, las correlaciones son muy débiles para D8 ($r=0,072$; sig. 0,738) y para D8 y 9 ($r=0,024$; sig. 0,912), e incluso negativas para la dimensión D9 ($r=-0,107$; sig. 0,618). El contraste para las diferencias medias entre las dimensiones de metacognición y la puntuación global en ECOS arroja correlaciones negativas, aunque sin llegar a ser significativas, siendo para la puntuación diferencia entre los totales pretest y posttest de la escala de $r=-0,151$, con una significación de 0,492. En cuanto al contraste entre las puntuaciones de la dimensión *Metaconocimiento*

estratégico en la prueba ECOS y las obtenidas para las dimensiones relativas a la metacognición en la escala, las correlaciones vuelven a ser negativas, mostrando la más débil, para la dimensión D8, un valor de $-0,064$ y una significación de $0,765$, mientras que la más fuerte, para la dimensión D9, mostró un valor de $-0,303$ y una significación de $0,149$.

Los contrastes entre las diferencias promedio pretest-postest de las dimensiones metacognitivas de la escala y de la puntuación global de ECOS ($N = 23$), arrojan valores negativos de escasa fuerza para D8, D9 y D8 y 9 (por ejemplo, D8 y 9 $r. -0,077$; sig. $0,728$), mientras que para la puntuación global, la correlación es positiva aunque carece de relevancia ($r. 0,004$; sig. $0,984$). Para el contraste con las diferencias medias obtenidas en la dimensión *Metaconocimiento estratégico* de la prueba ECOS, los valores obtenidos expresan también correlaciones negativas, pero esta vez de mayor fuerza. Para la dimensión D8 la correlación fue de $-0,387$, con una significación asociada de $0,068$; para D9 el valor es más débil, $r. -0,074$, sig. $0,737$; la correlación para la dimensión conjunta D8 y 9 llega a ser significativa, con un valor de $-0,414$ y una significación de $0,049$; y, por último, la correlación con la puntuación global de la escala fue de $-0,376$, con una significación de $0,077$. Estos resultados muestran paradójicamente que, a pesar de haberse producido mejoras tanto en las puntuaciones medias de la dimensión de *Metaconocimiento estratégico* de la prueba de comprensión, como en las dimensiones metacognitivas de la escala APSL, los cambios positivos observados no han conllevado una misma dirección. Ni las correlaciones entre las puntuaciones obtenidas entre estas dimensiones de ambas pruebas, ni entre las diferencias promedio pretest-postest, han resultado positivas, tal y como podía suponerse.

Por último, al igual que realizamos con la prueba de primaria, hemos efectuado el recuento de sujetos que obtuvieron ganancias en ambas pruebas, lo que puede ofrecernos una información valiosa acerca de la convergencia de los resultados. En este caso han sido 8 los alumnos que compartieron ganancias en ambas pruebas cuando se observa la puntuación global, y 9 si lo que observamos es el número de sujetos que obtuvo ganancias en la puntuación global de ECOS y en la dimensión conjunta D8 y 9.

Cuando hemos efectuado los contrastes adoptando el criterio de la obtención de un puntaje bajo en la prueba de comprensión, los resultados para los 11 sujetos de esta submuestra muestran igualmente una correlación negativa, aunque ligeramente mayor que para el total de la muestra, entre las diferencias medias para la puntuación global de ECOS y APSL ($r. -0,032$, sig. $0,925$). De igual forma se han comportado los datos de la correlación entre las puntuaciones diferencia para la puntuación global de la escala y las de la dimensión *Metaconocimiento estratégico* de ECOS; es decir, muestran un valor de correlación ligeramente superior al de la muestra completa, aunque sigue siendo marcadamente negativo ($r. -0,342$, sig. $0,304$). En cambio, para la dimensión D8 la correlación con la puntuación global de ECOS pasa de ser negativa a poseer un valor positivo ($r. 0,100$, sig. $0,770$), mientras que incrementa su valor negativo con la dimensión de *Metaconocimiento estratégico* de ECOS ($r. -0,481$, sig. $0,134$). Por su parte, la dimensión D9 acentúa su valor negativo en la correlación con la puntuación total de ECOS, mientras que suaviza su correlación negativa con la dimensión de la prueba de comprensión. Por último, la dimensión conjunta D8 y 9 muestra un incremento de su correlación negativa, tanto para la puntuación total de ECOS como

para la dimensión *Metaconocimiento estratégico* de esta prueba, correlación ésta última que llega a ser significativa ($r = -0,657$, sig. 0,028).

Aunque posteriormente profundizaremos en la significación y el análisis de estos resultados, podemos apuntar que, a tenor de los análisis efectuados para valorar la convergencia de las puntuaciones de mejora en las pruebas de comprensión y de ajuste personal-social, no estaríamos en condiciones de certificar unos valores altos de fiabilidad de estas mejoras observadas en ambas pruebas.

8.3 Resultados en el test de inteligencia general de Cattell

Tal y como apuntábamos en páginas anteriores, la aplicación del test de factor «G» de Cattell en nuestro estudio cumplía diversas funciones; por una parte, debía servir al propósito del control de la variable inteligencia entre los grupos experimental y control de la muestra; por otra, debía ser un elemento relevante a la hora de establecer algunos subgrupos de sujetos con dificultades de aprendizaje en la lectura dentro de la muestra, por ejemplo aquellos en los que, en función de la puntuación obtenida en la prueba de CI, baja o normal, podría determinarse la pertenencia del sujeto al subgrupo de *dificultades específicas en la decodificación* o bien al subgrupo de dificultades en la *comprensión global*. Si bien la consideración de estos criterios fue la que inicialmente nos condujo a la aplicación de la prueba, éstos no agotaban las posibilidades de su utilización en una investigación como la nuestra. De hecho, podríamos haber hecho uso de ella sólo en la fase pretest, habiendo obtenido de esta forma la información que nos permitía aplicar las anteriores premisas. Más allá de éstas, la aplicación de la prueba de Cattell en las fases pretest y postest del estudio, obedece también a la consideración de la posibilidad de obtener diferencias postratamiento en el desempeño de los sujetos experimentales, debidas al trabajo realizado en las sesiones de aplicación del dispositivo de evaluación dinámica EDPL.

Aunque en un principio esta posibilidad nos parecía remota, debido esencialmente a que el contenido del dispositivo aplicado no se relacionaba en absoluto con el que es evaluado en la prueba de CI, contemplamos posteriormente esta opción en función del siguiente criterio: la aplicación del dispositivo EDPL supone fundamentalmente un trabajo sobre procesos metacognitivos de conocimiento, control y regulación consciente del proceso de comprensión y de cada uno de los procesos implicados en la misma. Este trabajo se concreta en aspectos como: (a) la elicitación de una mayor reflexión de los sujetos sobre los propios procesos de acción encaminados a la solución de un determinado problema; (b) el modelado por parte del profesor/aplicador de formas de actuación guiadas por la reflexión y la regulación del proceso para la obtención de mejores resultados; y por último, (c) la aplicación de estrategias metacognitivas de comprensión, que incluyen fundamentalmente elementos de planificación del trabajo a realizar, procesos de supervisión constante de la actividad mientras ésta se está ejecutando, y autoevaluación del proceso.

A continuación nos disponemos a mostrar los resultados de los diversos contrastes realizados. Ya en el capítulo en el que analizábamos la composición y las peculiaridades de la muestra, hacíamos alusión a la equivalencia inicial de los grupos experimental y control, por lo que abordamos ahora la evaluación de las posibles diferencias postratamiento, tanto para los diversos subgrupos contemplados, como en función de los criterios formulados en el estudio para la valoración de los resultados.

(a) Criterio de designación previa del profesorado sobre las dificultades de aprendizaje

Comenzamos por tanto por presentar los resultados pretest y postest obtenidos en Cattell para los grupos experimental y control en los distintos subgrupos contemplados en la muestra en función de su grado de dificultad y de las condiciones de apoyo requeridas.

Tabla 59. Comparaciones pretest-postest en la prueba Cattell-2 para los distintos subgrupos contemplados en la muestra (DA = dificultades de aprendizaje; DIS = discapacidad; AI = recibe apoyo a la integración; DEP = privación socio-cultural; AC = recibe apoyo a ciclo o refuerzo educativo; AF = agrupamientos flexibles; RL = Refuerzo de Lengua; NA. = no recibe apoyo)

Prueba t para muestras relacionadas

Gr.	Grupos	N	Pretest	DT	Postest	DT	t	gl	Sign.
GE GC	Grupo completo	57 42	83,21 80,74	13,70 17,68	88,00 83,81	14,92 21,57	-3,360 -1,659	56 41	0,001** 0,105
GE GC	Grupo reducido (DIS, DEP o AI)	55 25	82,84 73,04	13,81 16,02	88,05 74,32	15,18 18,16	-3,612 -0,562	54 24	0,001** 0,579
GE GC	Primaria Grupo completo	30 10	81,43 78,70	14,30 16,67	86,50 73,90	16,06 19,93	-2,245 1,237	29 9	0,033* 0,247
GE GC	Secundaria Grupo completo	27 32	85,19 81,38	12,99 18,19	89,67 86,91	13,63 21,41	-2,622 -2,851	26 31	0,014* 0,008**
GE GC	Grupo DA	30 22	84,03 88,64	14,27 14,84	86,80 93,36	15,84 19,35	-1,416 -1,679	29 21	0,168 0,108
GE GC	Grupo Discapacidad	11 15	75,91 66,47	10,94 12,83	86,55 68,73	15,15 17,04	-3,482 -0,989	10 14	0,006** 0,339
GE GC	Grupo Deprivación	16 5	86,69 88,80	13,19 16,83	91,25 87,00	13,30 20,48	-1,746 0,262	15 4	0,101 0,807
GE GC	Grupo NA, AC, AF, RL	2 17	93,50 92,06	0,71 13,69	86,50 97,76	2,12 18,69	7,000 -1,848	1 16	0,090 0,083

La observación de los datos nos permite realizar una valoración muy positiva acerca de la significación de las mejoras obtenidas por el GE cuando han sido contabilizados todos sus efectivos. Las diferencias llegan a ser muy significativas estadísticamente para el GE, mientras que para el GC, aunque obtiene también mejoras postest, éstas no superan los valores estadísticos de rechazo de la hipótesis nula de igualdad de las medias. Estos resultados positivos vuelven a producirse con mayor magnitud cuando se realizan los contrastes con el grupo reducido de sujetos, es decir, cuando se toman sólo los efectivos que además de ser catalogados con dificultades de aprendizaje por parte de sus profesores u orientadores, presentan discapacidad, privación socio-cultural o reciben apoyo a la integración. Recuérdese que estos contrastes obedecían a ciertas desigualdades observadas en la conformación de los grupos experimental y control. Los resultados, como podemos ver, son muy positivos para el GE en contraposición a los obtenidos por el GC, que apenas incrementa ligeramente su puntuación media inicial. En este caso, es necesario puntualizar, no obstante, que el contraste realizado para examinar las posibles diferencias de partida en las puntuaciones de ambos grupos, da como resultado la no equivalencia de los grupos, con un valor t de 2,796 y una significación asociada de 0,007 para 78 grados de libertad. Este dato no invalida

lógicamente los resultados positivos del GE, pero sí introduce cautelas a la hora de establecer comparaciones con los resultados obtenidos por el GC.

Por otra parte, también los resultados positivos del GE se mantienen cuando la muestra se subdivide en las dos etapas de primaria y secundaria. Sobre todo, en la submuestra de primaria, las diferencias con respecto al GC son muy notables, ya que mientras que el grupo de tratamiento obtiene ganancias significativas en el postest, el GC muestra pérdidas que, en todo caso, no llegan a ser significativas. Cuando hemos analizado los grupos reducidos experimental y control para la submuestra de primaria, nos hemos encontrado que el número de sujetos del GE apenas varía, pasando de 30 a 28 sujetos, incrementándose las diferencias a favor del postest; por su parte, el GC no varía sus efectivos, obteniendo obviamente los mismos resultados. En cuanto a la submuestra de secundaria, observamos que el GE obtiene nuevamente mejoras significativas, pero en este caso también el GC muestra ganancias que llegan a ser muy significativas estadísticamente. Cuando hemos analizado los grupos reducidos para estos subgrupos, hemos obtenido idénticos resultados para el GE, ya que el número de sujetos que permanecía en el contraste era el mismo, y en cambio, resultados más modestos para el GC, que veía reducido el número de sujetos de 32 a 15, mostrando de todas formas mejoras significativas esta vez al nivel 0,05.

No llegan a ser significativas las mejoras obtenidas por el GE en el contraste con el alumnado que presenta DA en ausencia de discapacidad y deprivación socio-cultural. Como podemos observar, el incremento no resulta significativo, ni tampoco el que experimenta el GC, aunque su diferencia de medias es ligeramente mayor. Hemos reducido esta submuestra, seleccionando sólo aquellos sujetos que recibían AI. Para este nuevo subgrupo con DA que recibe AI los resultados indican ganancias para el GE un poco mayores, que superan los tres puntos, aunque no llegan a ser significativas estadísticamente ($t = -1,335$, sig. 0,196 para 21 grados de libertad). En cambio, para el GC, que sólo contaría con 5 sujetos en estas condiciones, los resultados indicaban una ligera ganancia postest de poco más de un punto, con una t observada de $-0,198$ y una significación de 0,853 para 3 gl.

Las ganancias más importantes se producen para el subgrupo de sujetos que presentaban discapacidad. Como podemos observar, las mejoras que obtiene el grupo de tratamiento son muy significativas, con más de 10 puntos de diferencia de promedio; mientras que el grupo control incrementa poco más de 2 puntos su puntuación promedio inicial. En este caso, las diferencias observadas en las puntuaciones iniciales de ambos grupos en este contraste, resta relevancia a la comparación directa entre las ganancias obtenidas, aunque deja intacta la significatividad de las mejoras obtenidas por el grupo experimental. Por su parte, el contraste para el alumnado con deprivación, en ausencia de discapacidad, ofrece resultados positivos para el GE, con ganancias postest cercanas a los cinco puntos, aunque no se muestran suficientes para rechazar la hipótesis de igualdad de diferencias. Por su parte, el GC experimenta ligeras pérdidas para este subgrupo, que no llegan a ser significativas. Como podemos observar, el número de sujetos del GC válidos para esta comparación es muy pequeño.

Por último, como aspecto destacable del último de los contrastes realizados, podemos indicar que, a pesar del exíguo número de sujetos del GE que cumplían las condiciones expresadas, es decir, no recibir apoyo de ningún tipo, o recibir sólo apoyo a ciclo, participar de agrupamientos flexibles, o recibir refuerzo educativo o refuerzo de

Lengua, todo ello en ausencia de discapacidad o deprivación socio-cultural, el contraste efectuado nos ha servido fundamentalmente para volver a comprobar precisamente la escasa presencia de sujetos del GE que presentaban estas características señaladas, mientras que el número de sujetos del GC es significativamente más alto. Otro aspecto importante, aunque con las limitaciones que supone esta comparación, es que los alumnos del GE que supuestamente se encuentran en mejores condiciones, no obtuvieron ganancias posttest, sino pérdidas, por lo que parece afianzarse la posibilidad de que el tratamiento haya tenido un efecto significativo precisamente en el alumnado que presentaba mayores dificultades. Profundizaremos en esta hipótesis a continuación, cuando abordemos los contrastes en función del segundo de los criterios formulados y contemplados en nuestro estudio.

(b) Criterio del profesorado + puntuación inicial en las pruebas ECO 1 y ECOS

Hemos querido conocer cuál ha sido el comportamiento en la prueba de CI de los sujetos que puntuaron inicialmente por debajo de una desviación típica de la media en las pruebas de comprensión, es decir, sujetos que obtuvieron una puntuación igual o inferior a 84 puntos. La conjunción de este criterio con el siguiente, formulado en torno a la obtención de bajas puntuaciones precisamente en la prueba de inteligencia, nos otorgará información valiosa acerca de si un determinado grupo de alumnos se ha beneficiado mayormente del trabajo realizado, y si ese grupo, como así ha venido ocurriendo en los anteriores contrastes para las pruebas de comprensión y actitudes, coincide con los sujetos que presentaban inicialmente un nivel de actuación más bajo. Hemos efectuado así mismo las comparaciones pretest-posttest para aquellos sujetos que obtuvieron puntuaciones por encima de 84 puntos y también para aquellos que obtuvieron puntuaciones «altas» en la fase pretest de las pruebas de comprensión. En este sentido, debemos destacar que no hemos hallado en nuestros análisis sujetos que obtuvieran puntuaciones por encima de una desviación típica de la media, es decir, igual o superior a 116 puntos; por tanto, hemos adoptado cualquier puntuación superior a la media, es decir, a los 100 puntos, para catalogar y seleccionar una puntuación como alta. Los resultados para estos contrastes se presentan en la siguiente tabla.

Tabla 60. Contraste pretest-posttest en Cattell en función de la designación previa del profesorado y de las puntuaciones obtenidas en las pruebas de comprensión

Gr.	Criterio seguido	N	Pretest	DT	Posttest	DT	t	gl	Sign.
GE	Criterio (a) +	24	79,33	13,96	86,42	16,43	-4,605	23	0,000**
GC	ECO 1/ECOS <= 84	15	76,33	17,30	72,67	19,70	1,406	14	0,181
GE	Criterio (a) +	33	86,03	13,00	89,15	13,85	-1,440	32	0,160
GC	ECO 1/ECOS > 84	27	83,19	17,73	90,00	20,33	-3,089	26	0,005**
GE	Criterio (a) +	6	83,50	12,88	89,00	13,25	-1,526	5	0,187
GC	ECO 1/ECOS > 100	5	97,60	13,79	99,40	17,39	-0,413	4	0,701

Podemos observar en el primero de los contrastes realizados que los resultados son muy positivos a favor del GE. Las diferencias entre las puntuaciones iniciales y las obtenidas con posterioridad a la aplicación del tratamiento son estadísticamente muy significativas. Estos datos parecen confirmar la tendencia observada en anteriores contrastes de que la aplicación del dispositivo de evaluación dinámica ha tenido un efecto positivo muy significativo para aquellos sujetos que presentaban inicialmente un

nivel más bajo. Por su parte, los resultados para el GC en este contraste indican pérdidas postest, aunque éstas no llegan a ser significativas. La relevancia de estos datos es grande, por cuanto de forma coherente con algunos de los estudios citados en este trabajo, serían precisamente aquellos sujetos que parten inicialmente de posiciones más bajas, quienes obtienen mejores resultados postratamiento.

El contraste en función de haber obtenido una puntuación mayor de 84 puntos arroja resultados positivos para el GE, que obtiene ganancias de más de 3 puntos en la media postest, aunque estas diferencias no llegan a ser significativas; sin embargo, en este caso, el GC obtiene mejoras muy significativas. Estos resultados parecen indicar que la intervención, aunque relevante, no ha tenido tanto peso en este subgrupo como en el de los sujetos que puntuaban bajo; en efecto, el hecho de que el GC haya obtenido mejoras significativas sin que haya mediado el tratamiento experimental, podría apuntar en esta dirección. Por último, el contraste para los sujetos que puntuaron por encima de 100 puntos muestra nuevamente ganancias importantes para el GE que, probablemente debido al exiguo número de sujetos en la comparación, no resultan significativas estadísticamente. En definitiva, podríamos concluir que el tratamiento habría tenido efectos positivos tanto para aquellos sujetos que puntuaron inicialmente bajo en las pruebas de comprensión, como para aquellos que obtuvieron mejores puntuaciones, siendo especialmente significativas las mejoras para los que partían de niveles de comprensión más bajos.

(c) Criterio del profesorado + puntuación inicial en las pruebas de comprensión + puntuación inicial en Cattell

El primero de los contrastes que se presenta a continuación tenía como objetivo conocer cómo evolucionaron las puntuaciones de aquellos sujetos que puntuaron bajo tanto en las pruebas de comprensión como en la propia prueba de inteligencia. Posteriormente comprobaremos las diferencias pretest-postest en esta prueba para aquellos que obtuvieron puntuaciones iniciales en CI que igualaban o se situaban por debajo de los 84 puntos. También efectuaremos comparaciones, al igual que hicimos anteriormente, para aquellos alumnos que presentaron mejores puntuaciones iniciales. A este respecto, tampoco en esta ocasión encontramos efectivos que hubieran superado, o al menos igualado, inicialmente los 116 puntos, por lo que optamos, como ya hicimos en el anterior contraste, por considerar puntuación «alta» la que superara la media, es decir, los 100 puntos. Como decíamos con anterioridad, los datos extraídos de estas comparaciones nos servirán para determinar los posibles efectos diferenciales de la aplicación del dispositivo en función del nivel cognitivo de los sujetos.

Los resultados obtenidos indican claras ganancias postest para el GE en el contraste realizado en función del tercero de los criterios formulados en nuestro estudio. El grupo de tratamiento ha ganado más de 8 puntos CI, mostrando ganancias postest muy significativas, mientras que el GC muestra pérdidas, que en todo caso no resultan significativas. Para el segundo de los contrastes expuestos, el que compara las puntuaciones pretest-postest para los sujetos que partían inicialmente con puntuaciones por debajo de una desviación típica de la media, también encontramos diferencias muy significativas a favor del postest en el GE; en cambio la evolución de las puntuaciones en el grupo no tratado apenas muestra pequeñas diferencias, que no resultan significativas estadísticamente.

Tabla 61. Contraste pretest-postest en Cattell en función de la designación previa del profesorado, las puntuaciones iniciales obtenidas en las pruebas de comprensión y la puntuación inicial en la prueba de CI

Gr.	Grupos y Criterio seguido	N	Pretest	DT	Postest	DT	t	gl	Sign.
GE	(a) + ECO 1/ ECOS	15	71,07	9,64	79,40	14,04	-4,115	14	0,001**
GC	<= 84 + CI <= 84	11	68,27	11,32	65,27	14,39	1,330	10	0,213
GE	CI <= 84	30	72,73	8,77	79,93	13,00	-3,741	29	0,001**
GC		26	69,35	11,03	71,69	14,76	-1,137	25	0,266
GE	CI > 84	27	94,85	7,16	96,96	11,50	-1,039	26	0,308
GC		16	99,25	7,83	103,50	15,50	-1,182	15	0,255
GE	CI > 100 (puntuadores altos)	7	104,71	2,43	102,00	16,45	0,476	6	0,651
GC		8	106,00	3,74	105,50	13,14	0,114	7	0,912

En cuanto a los contrastes efectuados con los sujetos que puntuaron inicialmente en Cattell por encima de los 84 puntos o por encima de los 100, los resultados indican ganancias postest en el primero de ellos para ambos grupos, aunque ninguna de las diferencias observadas resulta significativa. Por lo que respecta al segundo contraste, las diferencias encontradas son a favor del pretest en ambos grupos, mostrando ligeras pérdidas.

Hemos procedido a efectuar un análisis de regresión para las puntuaciones postest de la prueba de CI con el objetivo de evaluar qué porcentaje de la varianza era explicada por las puntuaciones iniciales obtenidas en la prueba. Para estos análisis decidimos utilizar el grupo experimental completo, es decir, coincidente con el primero de los criterios adoptados. En este caso, esta opción es la más conservadora, ya que las ganancias han sido mayores en función de los otros criterios formulados. Los resultados del análisis nos muestran un índice de correlación de Pearson de 0,72. La R^2 es de 0,519 y la R^2 corregida de 0,510. El modelo explica por tanto el 51% de la varianza observada en el postest, o dicho de otro modo, el pretest de la prueba CI explica el 51% de los resultados obtenidos por la muestra experimental en la misma prueba seis meses después. El análisis de varianza posterior arroja un valor F de 59,267 y una significación asociada de 0,000.

Los datos anteriores muestran la incidencia de las puntuaciones iniciales en los resultados postest. Por otra parte, aunque el análisis muestra la significatividad de esta relación, el porcentaje de varianza no explicada por el pretest es amplio, por lo que cabe esperar que la aplicación del dispositivo haya tenido un peso fundamental en las ganancias observadas. Estos análisis serán completados cuando se aborde la evaluación de la aplicación del dispositivo EDPL, a raíz del cual estableceremos también el valor de predicción de las distintas puntuaciones contempladas, así como del dispositivo con respecto a criterios externos como el rendimiento académico en el área de Lengua o la valoración del profesorado sobre el uso de los procesos implicados en la lectura.

Las DDP en Cattell

Recurrimos nuevamente al instrumento de tipo exploratorio que hemos usado para enriquecer los análisis estadísticos más psicométricos realizados. En este caso, también usaremos el grupo reducido de sujetos, tanto del GE como del GC, para efectuar los cálculos; como se recordará, se trata de aquellos alumnos con dificultades de

aprendizaje que precisaban apoyo a la integración o que presentaban también discapacidad o deprivación socio-cultural. Para proceder a este análisis, el primer paso que hemos dado ha sido calcular las diferencias pretest-postest, estableciendo una nueva puntuación que se concreta en el resultado de restar a la puntuación inicial la puntuación obtenida en el postest. Posteriormente hemos seguido todos los pasos apuntados anteriormente cuando abordamos este análisis para las pruebas de comprensión o ajuste personal-social. Con relación a los requisitos, ya expuestos con anterioridad, para la aplicación posterior de la prueba t, el uso de una prueba como Cattell garantiza en principio la distribución normal de las puntuaciones CI en la población. Igualmente, las medidas pretest y postest están realizadas en una escala de intervalos, con lo que las premisas se cumplen con claridad para la utilización del estadístico en estos contrastes. Los resultados obtenidos de aplicar la DDP a los datos de la prueba para el GE se presentan a continuación.

	Ganancias	Pérdidas	A favor de hipótesis %	En contra de hipótesis %
Sujetos	39	16	70,91	29,09
Puntos	411	124	76,82	23,18

Podemos observar que la diferencia en cuanto al número de sujetos que obtienen ganancias con respecto a aquellos que no mejoran sus puntuaciones, es notable. La traducción en porcentajes de esta diferencia muestra que más del 70% de los sujetos del GE obtuvieron ganancias postest en Cattell, lo que es ya un indicador positivo de la aplicación del dispositivo experimental. Estas diferencias entre ganancias y pérdidas aumentan cuando se considera la magnitud de las diferencias obtenidas por cada uno de los sujetos; en este caso, más del 76% de la evolución de los puntos CI ha sido positiva y favorable a la hipótesis de cambio. No obstante estos resultados positivos, que sin duda enriquecen el contraste psicométrico de medias pretest-postest, debemos constatar que la magnitud de las pérdidas también debe ser destacada, por cuanto cerca del 30% de los sujetos no obtuvieron mejoras postratamiento en esta prueba. Entre estos, 3 obtuvieron ganancias nulas y los 13 restantes obtuvieron pérdidas. Este dato también nos conduce hacia anteriores y posteriores reflexiones y análisis sobre el posible efecto diferencial del tratamiento en el grupo de sujetos que lo recibió. Como hemos visto, dentro del GE se distinguían por la fuerza de las diferencias obtenidas, aquellos sujetos que presentaban inicialmente mayores dificultades, tanto en la propia prueba de inteligencia como en la de comprensión de textos.

Hemos procedido a calcular la probabilidad estadística de que los resultados obtenidos se deban al azar. Para ello, como apuntábamos anteriormente, hemos usado la prueba t para muestras independientes. Los resultados los exponemos en la siguiente tabla.

*Tabla 62. Contraste en función de las DDP para el GE
Prueba t para muestras independientes*

	N	Diferencia media (pre — post)	DT	Prueba t	gl.	Sign.
Grupo experimental	39 (G) 16 (P)	10,538 7,937	6,532 7,009	-1,313	53	0,195

Los datos del contraste nos muestran la ausencia de diferencias estadísticamente significativas entre la diferencia promedio de las ganancias y la de las pérdidas; esto a pesar de que la ganancia postratamiento para el GE reducido era de 6 puntos. El dato obtenido no viene sino a confirmar la matización que este instrumento puede introducir en los resultados comparativos de medias globales. De esta forma, aunque obviamente no invalida la significatividad de las diferencias entre las puntuaciones pretest y postest, y con ello, la efectividad del tratamiento experimental, sí nos hace acentuar nuestra atención en una doble dirección; por una parte, en lo relativo a qué sujetos han obtenido ganancias, lo que ya ha comenzado a analizarse en páginas anteriores, y, por otra, referente a qué sujetos obtuvieron pérdidas o ganancia nula, así como por qué la magnitud de algunas pérdidas fue tan notable.

Hemos aplicado también los cálculos para la obtención de las DDP en el GC1, también reducido. Si bien en el GE el número de sujetos apenas se veía alterado, en el GC1 los sujetos que cumplían las condiciones expuestas eran sólo 25. De éstos, 14 integraron la columna de ganancias, lo que traducido en porcentajes representaba el 56% de la submuestra, y 11, de los cuales ninguno obtuvo ganancia nula entre pretest y postest, se alineaban con las pérdidas, lo que representaba el 44%. Los resultados del GC1 se muestran a continuación.

	Ganancias	Pérdidas	A favor de hipótesis %	En contra de hipótesis %
Sujetos	14	11	56	44
Puntos	125	93	57,34	42,66

*Tabla 63. Contraste en función de las DDP para el GC1
Prueba t para muestras independientes*

	N	Diferencia media (pre — post)	DT	Prueba t	gl.	Sign.
Grupo	14 (G)	8,928	6,390	-0,160	23	0,875
control	11 (P)	8,454	8,477			

Podemos observar que el GC también acumula un mayor porcentaje de sujetos a favor de hipótesis, aunque tanto este dato como el más completo que incorpora la magnitud de las diferencias obtenidas, denotan la práctica ausencia de diferencias relevantes. En el contraste con la prueba estadística, comprobamos que las diferencias promedio entre ganancias y pérdidas en este grupo no difieren prácticamente, lo que supone un valor muy pequeño del estadístico y una significación cercana a los valores que implican máxima igualdad entre las medias.

8.4 La Teoría de Respuesta al Ítem en las pruebas de comprensión de textos

La investigación que en los últimos años está teniendo lugar para mejorar el análisis de los cambios producidos como consecuencia de la aplicación de pruebas dinámicas, ha dado como resultado la aparición de algunos métodos novedosos que tratan de paliar las dificultades o limitaciones del análisis clásico. Entre estos métodos, quizás el más interesante y completo pueda ser el conocido como *Teoría de Respuesta al Ítem* (TRI) (Embretson y Reise, 2000; Hambleton, Swaminathan y Rogers, 1991). Básicamente, las

características principales de la teoría de respuesta al ítem pasan por considerar la valoración de la aptitud del sujeto en el contexto de su interacción con los ítems de la prueba; de este modo, las puntuaciones obtenidas varían en función de las respuestas del sujeto y de las propiedades del ítem. Uno de los principales valedores de este método fue Rasch (1980), quien ideó un modelo en el que la ganancia o pérdida para cada uno de los ítems de una prueba se predecía a raíz de la diferencia entre la dificultad del ítem y la aptitud del sujeto. De este modo, Rasch relacionó la dificultad de un ítem con la probabilidad con la que un determinado sujeto con un nivel determinado de aptitud se enfrenta a él (Sternberg y Grigorenko, 2003).

Nosotros hemos adoptado y adaptado este método en función de los datos obtenidos en las pruebas de comprensión. Nos ha parecido especialmente interesante aplicar algunas de las características elementales del método a los resultados de estas pruebas por dos razones: (a) tratarse de las pruebas directamente relacionadas con la evaluación de la comprensión de textos, objeto del estudio; y (b) haber realizado cálculos anteriores sobre la dificultad de los ítems de las pruebas ECO 1 y ECOS, obteniendo datos de interés con relación a las posibles diferencias que podrían observarse en función de los distintos niveles cognitivos presentes en los grupos.

Para efectuar estos análisis hemos operacionalizado cualitativamente los casos de la muestra en función de los resultados obtenidos en las pruebas. Esta clasificación la hemos adoptado de los estudios de Budoff y sus colaboradores (1987a; 1987b). Como se recordará cuando analizamos las propuestas de evaluación dinámica en el marco teórico de este trabajo, Budoff distinguía en sus investigaciones tres clases de sujetos: (a) aquellos que puntuaban inicialmente alto, es decir, obtenían puntuaciones altas antes de la fase de instrucción de la prueba dinámica, y que eran denominados *puntuadores altos*; (b) los *mejoradores*, que eran aquellos que obtenían bajas puntuaciones pretest, pero que conseguían ganancias significativas posttratamiento; y (c) los *no mejoradores*, que eran los que puntuaban inicialmente bajo y no conseguían mejoras con la instrucción. La consideración de esta clasificación cualitativa en nuestro estudio y su aplicación al método de análisis propuesto, ha estado lógicamente mediatizada por las características particulares del estudio, adaptando a las necesidades planteadas y al tipo de información buscada, las premisas y las propuestas realizadas originalmente.

Por lo que respecta a primaria, hemos encontrado 14 sujetos en el GE que obtuvieron mejoras postest, de los cuales 8 obtenían ganancias partiendo de puntuaciones iniciales por debajo de una desviación típica de la media (≤ 84). La media pretest para este subgrupo de sujetos *mejoradores* fue de 76,544 y la postest de 85,796, situándose por tanto la diferencia de medias en -9,252. El valor de la *t* fue de -3,351 y la significación asociada fue de 0,012 para 7 grados de libertad. Por su parte, 16 sujetos experimentales obtenían ganancia nula o pérdidas postest. De éstos, tan sólo 3 habían partido de puntuaciones bajas iniciales. La media pretest para este subgrupo de sujetos *no mejoradores* era de 80,936 y la postest de 78,693. Por último, el subgrupo de los *puntuadores altos* estaba integrado solamente por un alumno. Al igual que hicimos anteriormente, la consideración de una puntuación alta se hizo con aquellas que superaban la puntuación media, es decir, 100 puntos. El alumno obtuvo una media inicial de 109,346 y superó ésta en el postest, obteniendo 111,588 puntos.

Una vez extraídos estos tres subgrupos de la muestra experimental en primaria, la cuestión a la que intentamos dar respuesta consistía en averiguar en qué ítems había

conseguido sus ganancias el subgrupo de mejoradores. Para esto calculamos inicialmente los promedios de cada ítem para el subgrupo de los 8 sujetos que habían obtenido ganancias partiendo de posiciones iniciales bajas, y posteriormente contrastamos los promedios pretest con los posttest. Los resultados arrojaron 27 comparaciones favorables al posttest, es decir, que indicaban ganancias; 9 comparaciones que expresaban una ganancia nula y 7 comparaciones que indicaban pérdidas. Pero, ¿cuáles eran estos 27 ítems en los que los mejoradores habían obtenido ganancias? Revisamos los resultados obtenidos por la muestra piloto (N = 179) en cada uno de los ítems de la prueba ECO 1. Queríamos comprobar en qué ítems habían conseguido las mejoras los sujetos que hemos denominado, adoptando la terminología de Budoff y sus colegas, *mejoradores*. Supuestamente, serían los ítems más «fáciles» aquellos en los que los sujetos con menores puntuaciones iniciales habrían obtenido las ganancias posttest. Los investigadores de la TRI han puesto de manifiesto en los sucesivos modelos que se han desarrollado, el sesgo que podía suponer la consideración por igual de todos los ítems de una prueba a la hora de la cuantificación de los resultados, sin tener en cuenta tanto la dificultad de cada uno de ellos, como el nivel cognitivo de los sujetos que afrontaban la prueba. En este sentido, el análisis de una prueba de aptitud debería tomar en cuenta estas dos variables mencionadas. Nosotros hemos entresacado los ítems que resultaron «fáciles», es decir, con aciertos promedio por encima del 50%, y hemos relacionado éstos con aquellos en los que los mejoradores obtuvieron sus ganancias. También hemos tomado en cuenta los ítems «difíciles» o «muy difíciles», de forma que pudiéramos igualmente comprobar si los mejoradores dejaban de obtener mejoras posttest en los ítems más difíciles de la prueba o, por el contrario, mantenían sus ganancias también en éstos. En la tabla siguiente mostramos la evolución de las puntuaciones promedio para cada ítem de la prueba ECO 1.

Tabla 64. Evolución de las puntuaciones promedio de los ítems de ECO 1 y catalogación de dificultad para el subgrupo de mejoradores en función del promedio de aciertos de la muestra piloto (F = fácil, media > 0,50; D = difícil, media 0,25 - 0,50; MD = muy difícil, media < 0,25)

Ítems	Promedio pretest	Promedio posttest	Promedio de aciertos De la muestra piloto	Ítems	Promedio pretest	Promedio posttest	Promedio de aciertos de la muestra piloto
1	0,13	0,50	0,709 (F)	23	0,00	0,13	0,363 (D)
2	0,25	0,50	0,731 (F)	24	0,13	0,13	0,279 (D)
3	0,38	0,50	0,318 (D)	25	0,25	0,00	0,195 (MD)
4	0,13	0,38	0,441 (D)	26	0,00	0,13	0,474 (D)
5	0,13	0,25	0,469 (D)	27	0,00	0,25	0,312 (D)
6	0,25	0,25	0,296 (D)	28	0,00	0,13	0,396 (D)
7	0,13	0,50	0,379 (D)	29	0,00	0,13	0,597 (F)
8	0,13	0,13	0,223 (MD)	30	0,13	0,00	0,044 (MD)
9	0,13	0,00	0,223 (MD)	31	0,13	0,25	0,195 (MD)
10	0,00	0,50	0,402 (D)	32	0,00	0,13	0,134 (MD)
11	0,13	0,13	0,173 (MD)	33	0,00	0,13	0,536 (F)
12	0,13	0,00	0,173 (MD)	34	0,00	0,25	0,402 (D)
13	0,00	0,13	0,273 (D)	35	0,13	0,00	0,363 (D)
14	0,38	0,63	0,608 (F)	36	0,00	0,13	0,335 (D)
15	0,13	0,38	0,396 (D)	37	0,00	0,13	0,262 (D)
16	0,13	0,25	0,329 (D)	38	0,00	0,00	0,279 (D)
17	0,13	0,13	0,234 (MD)	39	0,13	0,13	0,217 (MD)
18	0,13	0,13	0,469 (D)	40	0,00	0,13	0,463 (D)
19	0,00	0,25	0,553 (F)	41	0,00	0,25	0,357 (D)
20	0,00	0,13	0,569 (F)	42	0,13	0,00	0,346 (D)
21	0,00	0,13	0,162 (MD)	43	0,13	0,13	0,134 (MD)
22	0,50	0,25	0,273 (D)				

Mediante la exploración de los datos podemos comprobar que en los 7 ítems que superaban el 50% de aciertos promedio para la muestra de aplicación piloto, obtuvieron ganancias los sujetos integrantes del subgrupo de mejoradores. No encontramos ningún ítem de los catalogados como «fáciles» en los que no consiguieran mejoras postest. Por otra parte, 17 ítems en los que mejora este subgrupo son ítems que fueron catalogados como «difíciles», con un promedio de aciertos entre 0,25 y 0,50. También obtuvieron mejoras en 3 ítems clasificados como «muy difíciles», cuyo promedio de aciertos para el grupo piloto se encontraba por debajo del 25%. En cuanto a las pérdidas —hemos contabilizado entre éstas las ganancias nulas, al no producirse un incremento en el postest, indicativo de mejora—, 9 de ellas se produjeron en ítems «muy difíciles», mientras que 7 se produjeron en los niveles de dificultad 0,25-0,50.

Por tanto, tras este análisis podríamos concluir que aquellos alumnos con bajo nivel inicial en la prueba de comprensión ECO 1, pero que mostraron ganancias significativas tras la aplicación del tratamiento experimental, obtuvieron mejoras postest en todos aquellos ítems que resultaron «fáciles» a la muestra piloto; además mostraron ganancias en otros ítems que habían resultado difíciles o muy difíciles al conjunto de sujetos al que se aplicó la prueba inicialmente —recuérdese que esta prueba podría haber resultado en exceso difícil para el alumnado de esta muestra piloto, con un índice medio de dificultad de 0,6487—. Por otra parte, el análisis de los datos también nos muestra que hasta en 5 ocasiones el subgrupo de mejoradores obtuvo puntuaciones postest que llegaron a superar a las del grupo piloto y en otras 8 igualaron o se acercaron mucho a éstas. Esto no debe hacernos olvidar que, a pesar de las importantes ganancias obtenidas por este subgrupo, su puntuación media postest se sitúa lejos aún de la obtenida por el grupo piloto. La media postest para los mejoradores fue de 85,796, con lo que superaban en más de 9 puntos la media obtenida en la fase inicial, al tiempo que superaban también los 84 puntos, punto de corte para establecer las puntuaciones consideradas como de bajo nivel; sin embargo, la puntuación del grupo piloto en la pasación postest de la prueba se situó en 106,185 puntos. Ante estos datos, debemos tener presente que el subgrupo de mejoradores estaba formado por aquellos alumnos que presentaban puntuaciones iniciales que se situaban significativamente por debajo de la media y que en su mayoría presenta dificultades importantes de aprendizaje asociadas en no pocas ocasiones a discapacidad. En síntesis, las mejoras de este subgrupo no se limitaron a los ítems a priori más fáciles de la prueba, sino que se ampliaron a otros en principio difíciles de resolver, incluso para sujetos sin ningún tipo de dificultades, como eran en su mayoría los que integraban el grupo de aplicación piloto.

¿Qué nos encontraríamos si contrastamos las medias pretest y postest del GE al completo tomando en cuenta sólo los ítems que a priori eran más fáciles? Efectuamos este contraste tomando en cuenta los 7 ítems que fueron catalogados como «fáciles» en función de los resultados del grupo piloto. Hemos realizado las comparaciones utilizando las puntuaciones directas en la prueba, es decir, sin la transformación a la que fueron sometidas. Esta puntuación es equivalente a la puntuación transformada. Los resultados para los 30 sujetos que integraron el GE completo mostraron la ausencia de significatividad en las diferencias, con una puntuación media inicial de 2,000 (DT = 1,531) y una puntuación postest de 2,067 (DT = 1,552). El valor de t observada fue de -0,186 y la significación asociada de 0,854 para 29 gl. Como se puede comprobar, el ligero incremento postest es similar al que se da para este mismo grupo con el total de ítems, por lo que, al menos para el GE al completo no parece haber tenido efecto la distinción del contenido de la prueba en función de la dificultad de los ítems que la

componen. En cambio, sí habíamos podido certificar anteriormente la relevancia que parece haber tenido este análisis cuando se toma en consideración también la capacidad y el desempeño de los sujetos que intervienen.

En educación secundaria hemos encontrado 21 sujetos en el GE que obtuvieron mejoras postest, de los cuales 11 obtenían ganancias partiendo de puntuaciones iniciales por debajo de una desviación típica de la media (≤ 84). La puntuación media en el pretest para este subgrupo de sujetos *mejoradores* fue de 79,387 y la postest de 88,988, resultando la diferencia de medias de -9,601. En el contraste de medias realizado, el valor de la *t* fue de -6,287 y la significación asociada de 0,000 para 10 grados de libertad. Por su parte, 4 sujetos de la muestra experimental obtenían ganancia nula o pérdidas postest. De éstos, tan sólo 1 había partido de puntuaciones bajas iniciales. La media pretest para este sujeto *no mejorador* fue de 79,519 y la postest de 78,072. Por último, el subgrupo de los *puntuadores altos* estuvo integrado por 5 sujetos. La puntuación media pretest en este caso se situaba en 105,271, mientras que la postest se iba a los 109,539 puntos, estableciéndose una diferencia de medias de -5,787.

Al igual que realizamos con la submuestra de primaria, tratamos de averiguar qué ítems habían ocasionado las ganancias para el subgrupo de mejoradores. El contraste de promedios pretest y postest de los ítems para este subgrupo nos ofrecía los datos siguientes: 41 comparaciones expresaban ganancias; 8 indicaban ganancia nula y 7 expresaban pérdidas postest. Procedimos posteriormente a revisar los resultados obtenidos por la muestra piloto, en este caso más reducida que en primaria ($N = 43$), en cada uno de los ítems de la prueba ECOS. Entresacamos los ítems que resultaron «fáciles», es decir, con aciertos promedio por encima del 50%, y relacionamos éstos con aquellos en los que los mejoradores obtuvieron sus ganancias. Igualmente tomamos en consideración los ítems «difíciles» o «muy difíciles», con la intención de cerciorarnos acerca de si los mejoradores obtenían también ganancias postest en los ítems más difíciles de la prueba. En la tabla siguiente mostramos la evolución de las puntuaciones promedio para cada ítem de la prueba ECOS.

Una exploración inicial de los datos de la tabla 65 establece una diferencia notable con los datos analizados anteriormente; se trata del mayor índice promedio de aciertos en la muestra piloto. Como pudimos comprobar cuando analizábamos el contenido de esta prueba en el capítulo dedicado a los instrumentos, el índice medio de dificultad para ECOS se situaba en 0,5125, dato a priori más adecuado globalmente para la valoración del nivel de comprensión en la muestra e indicativo de que la prueba ofrecía un nivel de dificultad asequible para la mayoría de los sujetos de la aplicación piloto. Al igual que hicimos anteriormente, hemos calculado el número de ítems «fáciles», es decir, con un promedio de aciertos en la muestra piloto que superaba el 50%, así como el número de los ítems «difíciles» (promedio de aciertos entre 0,25 y 0,50) y «muy difíciles» (promedio inferior a 0,25). Lógicamente, al ser el promedio global de aciertos superior en esta prueba, hemos encontrado menos ítems catalogados como «muy difíciles» y mayor número de los catalogados como «difíciles» o «fáciles». Podemos comprobar que de las 41 comparaciones que expresan ganancia postest, 18 lo hacen en los ítems «fáciles»; sin embargo, encontramos también 7 ítems «fáciles» en los que se obtienen pérdidas o ganancias nulas. Otras 21 comparaciones de mejora se obtienen en ítems catalogados como «difíciles» y las 2 restantes en ítems clasificados como «muy difíciles». Por lo que respecta a las pérdidas, entre las que hemos contabilizado las ganancias nulas al no producirse un incremento en el postest que fuera indicativo de

Tabla 65. Evolución de las puntuaciones promedio de los ítems de ECOS y catalogación de dificultad para el subgrupo de mejoradores en función del promedio de aciertos de la muestra piloto (*F* = fácil, media > 0,50; *D* = difícil, media 0,25 - 0,50; *MD* = muy difícil, media < 0,25)

Ítems	Promedio pretest	Promedio posttest	Promedio de aciertos De la muestra piloto	Ítems	Promedio pretest	Promedio posttest	Promedio de aciertos de la muestra piloto
1	0,18	0,27	0,791 (F)	29	0,00	0,18	0,558 (F)
2	0,09	0,09	0,558 (F)	30	0,00	0,18	0,628 (F)
3	0,45	0,18	0,581 (F)	31	0,00	0,36	0,535 (F)
4	0,27	0,27	0,395 (D)	32	0,00	0,27	0,651 (F)
5	0,00	0,09	0,558 (F)	33	0,00	0,27	0,488 (D)
6	0,55	0,73	0,767 (F)	34	0,09	0,27	0,279 (D)
7	0,36	0,45	0,674 (F)	35	0,09	0,36	0,674 (F)
8	0,55	0,27	0,884 (F)	36	0,00	0,09	0,349 (D)
9	0,55	0,36	0,744 (F)	37	0,09	0,18	0,256 (D)
10	0,18	0,36	0,372 (D)	38	0,00	0,18	0,256 (D)
11	0,09	0,18	0,814 (F)	39	0,00	0,27	0,419 (D)
12	0,18	0,36	0,674 (F)	40	0,00	0,36	0,581 (F)
13	0,00	0,09	0,488 (D)	41	0,09	0,27	0,302 (D)
14	0,09	0,09	0,256 (D)	42	0,09	0,27	0,488 (D)
15	0,36	0,55	0,744 (F)	43	0,00	0,09	0,372 (D)
16	0,55	0,73	0,814 (F)	44	0,00	0,00	0,512 (F)
17	0,27	0,09	0,651 (F)	45	0,09	0,36	0,558 (F)
18	0,36	0,45	0,953 (F)	46	0,00	0,09	0,488 (D)
19	0,00	0,27	0,442 (D)	47	0,00	0,18	0,419 (D)
20	0,27	0,27	0,395 (D)	48	0,00	0,27	0,488 (D)
21	0,09	0,00	0,163 (MD)	49	0,00	0,00	0,256 (D)
22	0,18	0,36	0,233 (MD)	50	0,00	0,18	0,395 (D)
23	0,45	0,27	0,535 (F)	51	0,00	0,27	0,349 (D)
24	0,09	0,45	0,279 (D)	52	0,00	0,00	0,279 (D)
25	0,09	0,00	0,419 (D)	53	0,09	0,09	0,116 (MD)
26	0,00	0,27	0,605 (F)	54	0,00	0,18	0,372 (D)
27	0,00	0,09	0,512 (F)	55	0,00	0,18	0,279 (D)
28	0,00	0,36	0,488 (D)	56	0,00	0,09	0,163 (MD)

mejora, 7 de ellas, como apuntamos anteriormente, se produjeron en ítems «fáciles», mientras que 6 se produjeron en los niveles de dificultad 0,25-0,50 y 2 en ítems «muy difíciles».

Otro dato de interés nos lo proporciona el hecho de que a pesar de los excelentes resultados obtenidos por el subgrupo de mejoradores, indicativos de mejoras sustanciales postratamiento, sólo en dos ocasiones sus puntuaciones posttest logran superar a las de la muestra piloto, y en tres más logran acercarse o igualarlas. Sin duda, este dato no nos hace olvidar que este subgrupo de alumnos partía con puntuaciones iniciales que se situaban significativamente por debajo de la media, y que sus ganancias, con ser muy significativas y relevantes desde el punto de vista estadístico y educativo, no pueden conducirles súbitamente hasta las puntuaciones obtenidas por sus compañeros que no presentan dificultades de lectura. Recordemos que la puntuación media posttest obtenida por este subgrupo se situó en 88,988, superando en más de 9 puntos la obtenida en el pretest, y superando así mismo la barrera de los 84 puntos, establecida como corte para determinar los sujetos que puntuaban significativamente más bajo; aún así, esta puntuación posttest se encuentra lejos de los 115,531 puntos conseguidos en la fase posttest por el grupo que conformó la muestra de aplicación piloto.

Entre las conclusiones de estos análisis podríamos destacar que la gran mayoría de los ítems catalogados como «fáciles» otorgaron ganancias en las comparaciones pretest-

postest para aquellos alumnos que partiendo de un nivel inicial bajo en la prueba de comprensión, mostraron ganancias significativas tras la aplicación del tratamiento experimental. Por otra parte, estos alumnos también mostraron ganancias en otros ítems que habían resultado difíciles o muy difíciles para el grupo de aplicación piloto. Al igual que ocurría con la submuestra de primaria, las ganancias postest de los mejoradores no se han limitado a los ítems que resultaron fáciles a la muestra piloto; por el contrario, la mayoría de los ítems cuyo contraste pretest-postest fue indicativo de mejoras, habían sido considerados como «difíciles» a tenor del promedio de aciertos en el grupo piloto.

Una última cuestión a contestar nos lleva a realizar los contrastes de puntuaciones medias pretest y postest para el GE tomando en consideración sólo los 25 ítems que habían resultado «fáciles» al grupo de aplicación piloto. La comparación, al igual que para el GE en primaria, la efectuamos para el grupo de tratamiento al completo y mediante la utilización de las puntuaciones directas obtenidas en la prueba. La puntuación media pretest para los 25 sujetos es de 8,360 (DT = 4,769) y la postest de 10,120 (DT = 5,403); el valor del estadístico t es de -2,097 y la significación de 0,047 para 24 grados de libertad. En este caso, observamos que el contraste es significativo al nivel 0,05, aunque debemos recordar que lo era a un nivel 0,01 cuando se tomaban en consideración todos los ítems de la prueba. Este dato, unido al que indicaba para el subgrupo de mejoradores una ganancia postest significativa igualmente al nivel 0,01, es coherente con el hecho de que muchos de los ítems en los que este subgrupo de mejoradores, sobre el que se realizaban los análisis anteriores, obtenía ganancias, eran elementos de la prueba a priori más difíciles, por lo que podemos confirmar que las mejoras observadas, tanto en este subgrupo como en el global del GE en esta etapa, se han obtenido tanto en los ítems a priori más asequibles, como en aquellos que resultaron más difíciles incluso para el grupo de aplicación piloto.

8.5 Los resultados de la muestra piloto en el postest

Como se recordará, una vez iniciado el estudio y en función de que las variaciones en el contenido de las pruebas de comprensión y de ajuste habían tenido ya lugar en su mayor parte, previamente a la aplicación piloto de las mismas, creímos oportuno volver a aplicar estas pruebas en la fase postest de nuestra investigación. El motivo principal consistía en recabar información sobre la evolución de los niveles de comprensión de textos y de ajuste personal-social en un grupo de alumnos en su mayoría sin dificultades de aprendizaje de la lectura. Esta información nos permitiría obtener un marco estándar de referencia con el que poder comparar los resultados obtenidos por los grupos experimental y control, no sólo sobre las puntuaciones iniciales en ambas pruebas, sino también sobre la evolución normativa que podría producirse sin intervención específica con sujetos que no presentan especiales dificultades. Supuestamente, como hemos apuntado en alguna ocasión a lo largo de este trabajo, los sujetos que no presentan problemas de aprendizaje evolucionarían positivamente de una forma «espontánea», es decir, sin necesidad de una intervención específica. En efecto, de no haber contado con este segundo grupo control (GC2), las puntuaciones obtenidas por algunos sujetos del GE en el postest carecerían de referencia externa alguna, a excepción del pretest del grupo piloto, de modo que si algún alumno superaba la puntuación inicial otorgada por la aplicación piloto, única referencia disponible en ese caso, desconoceríamos si esta ganancia le situaba cerca o a la par del grupo de sujetos sin dificultades, o, por el contrario, seguiría estando irremediablemente alejado. Hemos podido comprobar anteriormente cómo el conocimiento de las puntuaciones postest del grupo piloto en las

pruebas de comprensión nos ha servido para situar en contexto las ganancias del grupo de mejoradores.

Resultados posttest de la muestra piloto en las pruebas de comprensión

A continuación mostramos los resultados obtenidos por el GC2 en la prueba de comprensión ECO 1. Posteriormente completaremos el análisis que efectuábamos en el epígrafe anterior relativo a la TRI; los sujetos de este grupo, sobre todo aquellos sin dificultades de aprendizaje, deberían haber mostrado ganancias, tanto en los ítems a priori fáciles como en aquellos difíciles.

Tabla 66. Contraste pretest-posttest para el grupo de aplicación piloto (GC2) en la prueba de comprensión ECO 1

Prueba t de muestras relacionadas

Dimensiones	N	Pretest	DT	Posttest	DT	t	gl	Sign.
Puntuación total	171	100,453	14,382	106,185	15,675	-6,884	170	0,000**
Concepción de la lectura	171	100,532	15,838	102,059	16,199	-0,954	170	0,342
Metaconocimiento estratégico	171	100,775	14,601	105,588	15,813	-4,362	170	0,000**
Integración textual	171	100,409	15,559	105,338	15,677	-4,171	170	0,000**
Integración texto-conocimientos	171	100,313	14,568	104,720	16,709	-4,159	170	0,000**
Supervisión y autoevaluación	171	99,518	15,686	104,931	16,159	-3,725	170	0,000**

Podemos observar que salvo el contraste para la dimensión *Concepción de la lectura*, todas las diferencias resultan muy significativas. También podemos comprobar que las diferencias de promedio entre las fases inicial y final para este grupo piloto no son muy grandes, oscilando entre 4 y 5 puntos; recordemos que incluso para este grupo la prueba habría representado una dificultad notable. Junto a esta posible explicación, podría encontrarse un efecto *techo* para este grupo, al que, a pesar de que la prueba ofrezca la posibilidad de incrementar las puntuaciones pretest, precisamente debido a que contiene un buen número de ítems de dificultad media o alta, le habría podido resultar en exceso difícil remontar sus puntuaciones más allá de lo observado. Esta hipótesis hallaría cierto fundamento en los contrastes que realizábamos para el GE y para el GC1 en función de los distintos criterios observados, en los que obtuvimos diferencias de medias en algunos casos sensiblemente mayores a las observadas para la presente comparación.

Por otra parte, tal y como apuntábamos anteriormente, nos disponemos ahora a completar de alguna forma el análisis realizado en función de la Teoría de Respuesta al Ítem. Apuntábamos que este modelo venía a cubrir algunas de las limitaciones que el análisis estadístico clásico, basado en la comparación de medias grupales simples, podía tener. En este sentido, proponía la toma en consideración de la dificultad de los ítems componentes de una prueba, así como, en posteriores formulaciones, la de la capacidad

del sujeto que afronta la prueba. Fruto de nuestro interés por estos métodos de análisis, surgidos muy especialmente a la luz de la investigación sobre pruebas dinámicas, ha sido nuestra adaptación de estas ideas al contenido de nuestro estudio, con el fin de matizar y enriquecer el análisis. Hemos comprobado qué ítems resultaban «fáciles» o «difíciles» para el subgrupo de mejoradores, compuesto por sujetos del GE que obtuvieron bajas puntuaciones pretest y ganancias significativas postratamiento. Estos sujetos, a su vez, habían obtenido ganancias en todos los ítems considerados «fáciles» en función del grupo piloto, y, así mismo, habían obtenido también ganancias en muchos ítems catalogados como «difíciles» e incluso «muy difíciles». Queremos ahora comprobar dónde obtuvieron sus mejoras los sujetos de la muestra piloto en la prueba ECO 1. Como decíamos antes, estas ganancias deberían haberse producido con mayor intensidad precisamente en los ítems de mayor dificultad, ya que los ítems más fáciles habrían sido resueltos hipotéticamente sin problemas ya en el pretest. Los resultados nos muestran que para 38 ítems de los 43 que componen la prueba, las comparaciones indican ganancias postest. De estos ítems en los que se logró mejora, 8 habían sido catalogados como «muy difíciles» por encontrarse el promedio de aciertos en el pretest por debajo de 0,25; 23 fueron clasificados como «difíciles», con promedios de acierto inicial entre 0,25 y 0,50; y 7 fueron considerados «fáciles» a tenor de la puntuación promedio inicial, superior a 0,50. Debemos hacer constar que las mejoras observadas en los ítems de máxima dificultad fueron muy pequeñas en casi todas las ocasiones, lo que por otra parte resulta lógico dada la dificultad global que pareció mostrar la prueba. Así mismo, las mejoras se produjeron en todos los ítems considerados «fáciles», que como podemos comprobar fueron muy pocos. En cuanto a las pérdidas, éstas tuvieron lugar en los 4 ítems restantes catalogados como «muy difíciles» y en uno clasificado como «difícil». Por tanto, a tenor de estos datos, podríamos afirmar que este grupo piloto, compuesto en su mayor parte por alumnado que no presentaba dificultades de aprendizaje, obtuvo sus mejoras precisamente de los ítems más difíciles, con un 81,58% de ítems de mejora catalogados como difíciles o muy difíciles. A pesar de esto, como apuntamos anteriormente, las diferencias de promedio en estos ítems difíciles fue mayormente pequeña y los mayores incrementos postest fueron obtenidos en los ítems con promedio de aciertos inicial superior al 50%. Las ganancias de este grupo tuvieron lugar de forma generalizada, incluyendo la mayoría de los ítems difíciles y muy difíciles, pero también los que representaron inicialmente un menor nivel de dificultad.

Nos adentramos ahora en los resultados obtenidos por el grupo piloto en secundaria. Al igual que hemos hecho con la prueba de primaria, analizaremos aquí también la evolución de las puntuaciones, así como los ítems en los que este grupo obtuvo sus ganancias postest. En este sentido, la exploración de los datos nos ofrece nuevamente la evidencia de que el grupo con el que se realizó la aplicación piloto, obtuvo un progreso significativo en su nivel de comprensión de textos a lo largo del curso escolar en el que tuvo lugar la fase experimental del estudio. De este modo, los resultados observados, al igual que los obtenidos en primaria, eran los esperados para este grupo, formado en su mayoría, como hemos remarcado, por alumnado que no presentaba ningún tipo de dificultad relacionada con la lectura. En condiciones normales, es esperable que los sujetos sin dificultades progresen en sus niveles de comprensión de textos durante un curso escolar; si no fuera así, estaríamos probablemente ante uno de los siguientes supuestos: (a) casos encubiertos de dificultades relacionadas con la comprensión; (b) deficiencias en los procesos de enseñanza; o (c) falta de sensibilidad e idoneidad de la prueba para captar los cambios producidos. Por el contrario, observamos ganancias postest para todas las comparaciones efectuadas, y diferencias muy significativas para la

Tabla 67. Contraste pretest-postest para el grupo de aplicación piloto (GC2) en la prueba de comprensión ECOS

Prueba t de muestras relacionadas

Dimensiones	N	Pretest	DT	Postest	DT	t	gl	Sign.
Puntuación total	37	108,845	14,095	115,530	15,200	-4,102	36	0,000**
Concepción de la lectura	37	104,289	17,370	108,972	17,291	-1,363	36	0,181
Metaconocimiento estratégico	37	105,448	16,692	112,272	14,653	-3,031	36	0,005**
Conocimientos previos	37	105,969	13,206	113,872	16,039	-3,321	36	0,002**
Integración textual	37	108,927	16,247	111,789	16,224	-1,562	36	0,127
Integración texto-conocimientos	37	110,551	14,821	113,834	14,195	-1,529	36	0,135
Supervisión y autoevaluación	37	102,856	16,543	115,895	16,833	-3,802	36	0,001**

puntuación global, así como para las dimensiones *Metaconocimiento estratégico*, *Conocimientos previos* y *Supervisión y autoevaluación*. Es destacable que las dos dimensiones más directamente relacionadas con los procesos metacognitivos obtengan las mayores diferencias de medias y ganancias muy significativas. Así mismo, podemos observar que dos de las dimensiones que no obtuvieron diferencias significativas consiguieron puntuaciones pretest más altas que el resto de dimensiones. En este sentido, aunque esto nos podría conducir a la consideración de un cierto efecto *techo* para este grupo en esta prueba, al menos en las dimensiones citadas; el hecho de que los ítems de la prueba se distribuyeran en un gradiente de dificultad suficiente para admitir puntuaciones que se tradujeran en mayores ganancias postest; así como las propias mejoras observadas, cercanas a la significatividad estadística, no parecen apoyar esta posibilidad.

En cuanto al análisis de los ítems ganadores en esta prueba para el grupo piloto, hemos obtenido los promedios de aciertos pretest y postest y los hemos comparado. Ya comprobamos anteriormente qué elementos de la prueba resultaban «fáciles» o «difíciles» para el subgrupo de mejoradores, formado por alumnado que obtuvo puntuaciones pretest por debajo de una desviación típica de la media y ganancias significativas postratamiento. Estos sujetos habían obtenido ganancias en 41 ítems; entre ellos, la mayoría de los ítems considerados «fáciles» en función del grupo piloto, así como también en muchos ítems catalogados como «difíciles» e incluso «muy difíciles». Nos disponemos ahora a comprobar dónde obtuvieron ganancias los sujetos de la muestra piloto en ECOS. Es de alguna forma esperable que estas ganancias se produzcan con una mayor intensidad en los ítems de mayor dificultad, al considerarse que los ítems más fáciles de la prueba habrían sido ya resueltos sin problemas en el pretest. De encontramos claramente con estos resultados —lo que ya hemos comprobado en parte cuando analizamos los resultados para el subgrupo de mejoradores—, estaríamos confirmando una de las críticas más severas al uso de procedimientos clásicos de evaluación para abordar la investigación con pruebas

dinámicas, y reforzaría a su vez la relevancia de efectuar los análisis complementarios que abordaremos en el siguiente capítulo, así como el análisis basado en la TRI que hemos adoptado en éste.

Los resultados nos muestran 41 comparaciones que indican ganancias posttest, es decir, el mismo número de ganancias que las obtenidas por el subgrupo de mejoradores en esta misma prueba. De estos 41 contrastes ganadores, 16 mejoras fueron obtenidas en ítems «fáciles», es decir, con promedios de acierto en el pretest que superaban 0,50; 19 fueron obtenidas en ítems catalogados como «difíciles», con promedio inicial de aciertos entre 0,25 y 0,50; y por último, 6 fueron conseguidas en ítems «muy difíciles», con promedio inicial de aciertos inferior a 0,25. Por otra parte, 8 de las pérdidas fueron obtenidas en ítems «fáciles» y 3 en ítems «difíciles», mientras que 4 contrastes manifestaban ganancias nulas. Hemos constatado al explorar los resultados que las pérdidas son casi en su totalidad de fuerza mínima, produciéndose en la mayoría de los casos en ítems en los que la puntuación pretest era ya bastante alta. Por otra parte, también es destacable que se produzcan hasta 12 contrastes significativos estadísticamente a favor del posttest, lo que da una idea de la fuerza de las ganancias, sobre todo cuando han sido obtenidas partiendo de puntuaciones iniciales más bajas. En este sentido, podemos concluir que este grupo piloto también parece haber respondido sin fisuras a la hipótesis que planteábamos anteriormente, por la cual los sujetos sin dificultades expresarían principalmente sus mejoras en aquellos ítems que resultaron más difíciles en la fase pretest, mostrando menos potencia en los contrastes de los ítems que ya fueron fáciles inicialmente. En números, el 61% de los contrastes ganadores fue obtenido con ítems difíciles o muy difíciles.

Resultados posttest de la muestra piloto en la escala de ajuste personal-social

La aplicación piloto de la escala de ajuste personal-social APSL se realizó en dos grupos clase de primaria (5º y 6º) y otros dos grupos de secundaria (2º y 4º). Como ya apuntamos en páginas anteriores, al igual que ocurriría con la pruebas de comprensión, una vez finalizada la aplicación piloto creímos oportuno volver a pasar la escala al mismo grupo de sujetos en el posttest, ya que las variaciones sufridas por la escala habían tenido lugar sobre todo tras la validación de expertos y hacían referencia a elementos que no afectaban directamente al contenido de los ítems, tales como la corrección de la prueba, la readscripción de algún ítem o la interpretación cualitativa y cuantitativa de la escala. Las razones para evaluar la evolución de las puntuaciones en este segundo grupo control son semejantes a las ya expresadas para las pruebas de comprensión, destacando el interés que guardaba para nosotros la comprobación de la evolución de procesos relativos a las actitudes, la motivación o las expectativas relacionadas con la lectura cuando estos aspectos no serían trabajados —al menos de forma controlada— de forma específica en los grupos piloto. Ya hablamos anteriormente de las dificultades que entrañaría un cambio significativo de actitudes, necesitado al parecer de una intervención dirigida de forma específica a tal fin y de mayor duración que la acometida por nosotros; sin embargo, este cambio se había producido en el grupo experimental cuando la conformación del mismo se efectuaba en función de los criterios segundo y tercero, es decir, precisamente cuando el nivel de ajuste personal-social de los sujetos en la fase inicial era significativamente más bajo, y cuando este nivel de ajuste bajo se veía acompañado también por un bajo rendimiento cognitivo. En estos casos, las mejoras del GE habían resultado ser significativas desde la óptica estadística, con lo que la aplicación del dispositivo EDPL habría supuesto un

trabajo efectivo en el difícil terreno de las actitudes relacionadas con la lectura para aquel grupo de sujetos que más dificultades presentaba en este campo.

Nos disponemos ahora a comprobar cómo han evolucionado las puntuaciones del grupo piloto; recordamos que al tratarse de grupos-clase, hablamos en su mayoría de sujetos sin dificultades de aprendizaje de la lectura. Ya hemos visto, tal y como esperábamos, que la evolución de las puntuaciones en las pruebas de comprensión indicaba ganancias muy significativas para el grupo piloto (los cuatro grupos que componen esta muestra piloto para la escala APSL, también formaron parte del grupo que participó en el pilotaje de las pruebas de comprensión). En este caso, como hemos apuntado, la hipótesis es más modesta con respecto a las ganancias que puedan obtenerse. Las razones estarían fundamentalmente en la falta de intervención específica y en las dificultades intrínsecas que conlleva el cambio de actitudes; por otra parte, debíamos contemplar la posibilidad de que las puntuaciones iniciales expresaran ya un nivel alto de ajuste, que difícilmente pudiera ser rebasado en el postest; posibilidad ésta indicativa de un posible efecto *techo*; no obstante, esta posibilidad parecía tener aquí menos fuerza, ya que las puntuaciones iniciales, aunque superaban significativamente a la obtenida por el alumnado que conformaba los grupos experimental y control, no eran excesivamente altas.

Tabla 68. Contraste pretest-postest para el grupo de aplicación piloto (GC2) en la escala de ajuste personal-social APSL

Prueba t de muestras relacionadas

Dimensiones	N	Pretest	DT	Postest	DT	T	gl	Sign.
Puntuación global	87	104,460	15,644	103,728	17,011	0,527	86	0,600
D1 Conocimiento de las propias limitaciones...	87	105,232	15,682	104,456	16,195	0,503	86	0,616
D2 Conocimiento de las propias expectativas...	87	105,162	15,036	103,900	15,436	0,844	86	0,401
D3 Conocimiento de la propia actitud...	87	104,333	16,060	101,810	16,672	1,564	86	0,122
D4 Conocimiento de los propios intereses...	87	101,144	16,003	101,079	17,447	0,043	86	0,966
D5 Percepción del apoyo del contexto...	87	100,591	14,753	101,448	16,152	-0,457	86	0,649
D6 Percepción de la relevancia del tema...	87	102,595	13,570	104,385	14,214	-1,296	86	0,198
D7 Conoc. de propias estrategias de relación...	87	101,377	15,392	100,819	16,179	0,313	86	0,755
D8 Metaconocimiento estratégico de lectura	87	103,557	16,249	102,677	16,678	0,510	86	0,611
D9 Reconocimiento de cap. de generalización...	87	100,641	16,749	102,634	14,448	-0,905	86	0,368
Dimensiones 8 y 9	87	103,375	16,094	103,288	16,268	0,051	86	0,960

Lo primero que constatamos al explorar los datos es la ausencia de diferencias significativas en los contrastes efectuados. Ni la puntuación global de la escala, ni las distintas dimensiones contempladas, mejoran significativamente en el postest; es más, se produce una ligera pérdida que, en cualquier caso, carece de significatividad. Por el contrario, se observa una gran estabilidad en las puntuaciones obtenidas en las distintas fases del estudio, lo que parece confirmar de manera fehaciente lo apuntado líneas más arriba relativo a la resistencia al cambio que plantean las actitudes, los intereses o las expectativas, así como a la necesidad de intervención específica para provocar un cambio significativo en este campo. Posteriormente, cuando abordemos el análisis y la discusión de estos resultados, profundizaremos en la posibilidad de la existencia de un efecto *techo* en las puntuaciones obtenidas por este grupo.

Hemos querido comprobar si la ausencia de diferencias observadas entre pretest y postest se mantiene para los contrastes en función de la etapa educativa. En este sentido, hemos hallado que tanto para primaria como para secundaria las puntuaciones obtenidas mantienen como característica principal la igualdad entre las dos fases. Sin embargo, las puntuaciones en primaria son sensiblemente mayores que en la etapa superior. Para los 52 alumnos de primaria, la media pretest en la puntuación global se situó en 107,825 (DT = 15,444) y la postest en 107,145 (DT = 17,597), con un valor t de 0,383 y una significación de 0,703 para 51 grados de libertad. En cuanto a las distintas dimensiones contempladas, todos los contrastes realizados indican una igualdad similar entre las puntuaciones. Como aspecto a destacar, la puntuación significativamente más baja obtenida por este grupo en la dimensión D9, *reconocimiento de la capacidad de transferencia y generalización*, con relación al resto de dimensiones; con pretest y postest en torno a los 100 puntos. Por lo que respecta a secundaria, los 35 alumnos que realizaron las pruebas en pretest y postest obtuvieron una media inicial de 99,460 (DT = 14,771) y una postest de 98,652 (DT = 14,936), con un valor t observado de 0,358 y una significación asociada de 0,723 para 34 gl. Como aspectos destacables en referencia a los contrastes para las dimensiones, la D3, *conocimiento de la propia actitud ante la lectura*, obtiene pérdidas casi significativas, con una diferencia de medias de 4,238 puntos a favor del pretest y un valor de significación de 0,059. La dimensión D4 por su parte, *conocimiento de los propios intereses y motivaciones*, obtiene valores pretest y postest significativamente más bajos que el resto de dimensiones, con 92,286 puntos iniciales y 91,480 finales. Por último, y al contrario de lo que ocurría en primaria, la dimensión D9 obtiene las mayores ganancias postest, pasando de 101,474 puntos en el pretest a 105,808 en el postest, con un valor t de -1,268 y una significación de 0,213.

9. Evaluación complementaria

La evaluación complementaria: esquema del análisis propuesto

Cuando abordábamos en la primera parte de este trabajo el análisis de los modelos de comprobación dinámica, lo hacíamos desde una doble perspectiva; por una parte, tratábamos de conocer en profundidad las *bases teóricas y metodológicas* que sustentaban y daban coherencia al modelo, adentrándonos en aspectos relativos a la teoría de base, el formato de las pruebas propuestas, la naturaleza de los materiales de las mismas, las poblaciones a las que se dirigían, las medidas resultantes o la metodología de aplicación; por otra parte, hacíamos un recorrido por los estudios a los que habían dado lugar, realizando una síntesis de la *investigación* y una *revisión crítica del modelo*, donde prestábamos especial atención a elementos como el valor predictivo de las pruebas, la posibilidad de generalizar los resultados en función de la validez ecológica de los mismos, la validez y fiabilidad en general de que hacían gala, y, en síntesis, a todos aquellos aspectos que nos pudieran servir para valorar la idoneidad de la propuesta. Del mismo modo, cuando analizamos el contenido del dispositivo de evaluación dinámica EDPL en el capítulo dedicado a los instrumentos, ya hicimos referencia a aspectos que podrían englobarse dentro de las bases teóricas y metodológicas del modelo propuesto, como la población a la que se dirige, el formato y el contenido de las actividades propuestas, la naturaleza de la metodología de aplicación o los objetivos que fundamentalmente se pretenden con su utilización en el contexto escolar. Nos disponemos por tanto en este capítulo a dirigir nuestra atención hacia elementos que nos permitan realizar una valoración acerca de su idoneidad y validez, tanto desde un punto de vista psicométrico como ecológico. En definitiva, se trata de escribir ahora la «segunda parte» de nuestro análisis sobre el dispositivo experimental, es decir, exponer una *revisión crítica del modelo* EDPL.

Hemos adoptado para la evaluación de la aplicación del dispositivo de evaluación dinámica y, especialmente, de los datos empíricos, algunos elementos del *perfil de cuatro puntos* propuesto por Sternberg y Grigorenko (2003). El primero de estos puntos hace referencia a la *información comparativa*, en el que se examinan fundamentalmente las características diferenciadoras del modelo, la calidad de los datos obtenidos y el tipo de información que ofrece. La diferenciación se establece básicamente con relación a las pruebas convencionales que evalúan el mismo constructo y se pretende comprobar si la propuesta metodológica realiza aportaciones relevantes a la información que puede obtenerse con la pasación de una prueba clásica. En segundo lugar, se valora el *poder de predicción*, a través del cual se quiere obtener información sobre hasta qué punto la propuesta metodológica podría otorgar datos que faciliten una mejor predicción del rendimiento del sujeto y, como consecuencia, una intervención posterior mejor orientada a la optimización del proceso. El tercer punto hace referencia al *grado de eficacia* con el que la propuesta se presenta, en función del tiempo y del esfuerzo que son necesarios invertir en su aplicación para la obtención de la información deseada. Este grado de eficacia de la prueba dinámica se compara con el de pruebas estáticas y se valora hasta qué punto exige un mayor esfuerzo y tiempo de aplicación en función de la cualidad de la información recogida y de las propensiones que puedan realizarse de cara a la mejora de la acción educativa. Por último, el cuarto punto contemplado por

Sternberg y Grigorenko se refiere a la *solidez de los resultados*. Este punto se dirige a valorar si los resultados obtenidos en la evaluación de un dispositivo de evaluación dinámica han sido replicados por otros grupos de investigación o gozan de algún otro respaldo científico.

Como podrá deducirse, dada la novedad de la propuesta que realizamos, algunos de estos puntos, en parte o en su totalidad, no podrán ser abordados en nuestra revisión. Tal es el caso del cuarto punto, ya que hasta ahora sólo podríamos ofrecer datos empíricos o referencias teóricas de base que, en todo caso, avalarían la construcción de un dispositivo de estas características, pero no disponemos lógicamente de datos que revelen la replicación de nuestras conclusiones por parte de otros grupos de investigación. En cuanto a los datos empíricos y las referencias de base que aconsejarían y avalarían la construcción de un dispositivo que abordase la evaluación dinámica de procesos implicados en la lectura en sujetos que presentan especiales dificultades en la lectura, el capítulo dedicado a la valoración de los distintos modelos de comprobación dinámica, así como el que contiene las bases teóricas y metodológicas de nuestra propuesta, exponen sobrados datos de investigación y argumentos sólidos sobre la conveniencia de plantear este tipo de propuestas metodológicas.

Por otra parte, además de guiarnos en un principio por el perfil de los distintos puntos expuestos, profundizaremos y nos detendremos igualmente en algunos aspectos de nuestra investigación que pensamos han tenido —o pueden llegar a tener— una especial relevancia en su desarrollo o en los resultados obtenidos. Tal es el caso de la valoración cualitativa del proceso de aplicación por parte de los profesionales colaboradores, la evaluación del dispositivo por parte de estos mismos profesionales, o la re-elaboración, a raíz de la aplicación experimental del dispositivo, de una *matriz de patrones metodológicos de evaluación dinámica de la lectura*.

Por último, en este capítulo también nos adentraremos en la valoración de los cuestionarios sobre expectativas e ideas relacionadas con las dificultades de aprendizaje de la lectura. Ya expusimos cuando abordamos los distintos instrumentos del estudio un pequeño análisis de sus contenidos, así como de los posibles objetivos que podían guiar la valoración de sus resultados. Trataremos por tanto de enriquecer la información que nuestra investigación puede ofrecer sobre la evaluación de la lectura en sujetos con dificultades de aprendizaje, con el análisis del pensamiento de padres y profesores acerca de las posibilidades concebidas a este tipo de alumnado, así como acerca de las acciones emprendidas para favorecer la mejora de su rendimiento.

9.1 La evaluación de la aplicación del dispositivo de evaluación dinámica de procesos implicados en la lectura (EDPL)

Tal y como subraya Minick (1987), una característica fundamental y compartida por todos los modelos dinámicos de evaluación, es la hipótesis, a menudo transformada en convicción, de que la acción cognitiva en colaboración presenta mayores posibilidades de obtención de información valiosa para determinar el potencial de aprendizaje de un sujeto. Nuestra propuesta también comparte esta convicción y la presente investigación no es sino un intento más de ponerlo de manifiesto empíricamente, esta vez con relación a los procesos de lectura y comprensión.

Para acercarnos al logro del objetivo expresado en las líneas anteriores, además de tratar de obtener la mayor y más idónea información posible acerca de la aplicación experimental del dispositivo EDPL a través de los contrastes estadísticos de las puntuaciones medias de los diversos grupos contemplados en el estudio, hemos considerado relevante al mismo nivel la valoración cualitativa efectuada sobre el proceso de aplicación del dispositivo, sobre las dificultades encontradas en su puesta en práctica, así como sobre los resultados obtenidos. Igualmente, se ha considerado clave en nuestra investigación la determinación de una serie de patrones metodológicos utilizados por los aplicadores durante el desarrollo del proceso de mediación, que podrían explicar las mejoras observadas en algunos sujetos. Este análisis más cualitativo de los datos se relaciona mejor con algunos de los objetivos más relevantes de nuestra investigación, expresados a lo largo de este trabajo; en efecto, la consideración de la eficacia de un dispositivo de evaluación dinámica que pretende ofrecer información útil sobre los procesos implicados en la comprensión en sujetos con especiales dificultades de aprendizaje de la lectura, no puede descansar exclusivamente sobre la obtención de ganancias medias por parte de un grupo entre las fases inicial y final del estudio; esta información sería extremadamente pobre y carente de *significación práctica* (Botella y Barriopedro, 1995; Bickman y Rog, 1998), es decir, dejaría de dar respuesta a la principal finalidad de toda evaluación educativa: estar orientada a la acción, el cambio y la mejora (Lizasoain y Joaristi, 2000). La pretensión de dar algunas respuestas a las cuestiones que los educadores considerarían cruciales con relación a la evaluación de los procesos de comprensión, a la detección de las dificultades concretas de aprendizaje o a la determinación de vías de solución, pasa inexorablemente por analizar el *proceso de aprendizaje* de los sujetos que participaban del grupo experimental al que se aplicó el dispositivo, y también el *proceso de enseñanza/mediación* seguido por los profesores aplicadores durante el desarrollo de las sesiones. Esta evaluación del proceso incluye también referencias a categorías e indicadores difícilmente cuantificables, pero que pueden llegar a ser de gran utilidad en la práctica educativa y en la evaluación psicopedagógica.

La lógica seguida a lo largo del desarrollo del capítulo anterior, en el que además de presentar los datos de los contrastes realizados, buscábamos ofrecer algunas pistas que guiaran la posterior discusión sobre los mismos, apuntando posibles análisis cualitativos sobre los datos cuantitativos obtenidos; y, muy especialmente, la lógica seguida durante el análisis que ofrecemos en este capítulo, ha sido la de establecer un proceso flexible e interactivo, tratando, por una parte, de tener presente y de dar respuesta a las cuestiones e hipótesis formuladas en nuestro estudio previamente, así como, por otra, utilizando técnicas e instrumentos, tanto cuantitativos como cualitativos, para la obtención de la información deseada. A continuación, y antes de adentrarnos en el análisis efectuado en función del perfil de cuatro puntos, exponemos algunos aspectos relacionados con elementos y decisiones fundamentales de nuestro estudio, que nos permitirán posteriormente acometer el análisis de los resultados obtenidos.

Control experimental: análisis de las dificultades y vías de solución

Para comenzar, lo primero que debemos poner de manifiesto, tal y como ya apuntamos con anterioridad, es que a pesar del control ejercido desde el Grupo de Trabajo en el que participaron todos los profesores aplicadores; control sobre el contenido y la metodología de aplicación, así como sobre la temporalización de las actividades que se debían llevar a cabo, la puesta en práctica experimental de EDPL ha mostrado

importantes diferencias en los aspectos citados en función de los profesionales que realizaron la aplicación. Estas diferencias, lógicamente, añaden ciertas dificultades al análisis sobre el proceso de aplicación y sobre la valoración de los resultados obtenidos. En efecto, el número de actividades realizadas no fue finalmente el mismo para todos los aplicadores, así como no todos los profesores realizaron las mismas tareas. Ya comentábamos que, en parte, este hecho era inevitable y había sido ya tenido en cuenta por nosotros de forma previa; por otra parte, determinadas circunstancias acaecidas durante el desarrollo de la fase experimental, ajenas a nuestro estudio pero consustanciales a la investigación en el contexto escolar, han afectado en forma desigual a los grupos de aplicación y se han visto reflejadas de una u otra forma en los análisis y en los resultados obtenidos.

Recogemos a continuación en una tabla las actividades realizadas por cada aplicador, así como los procesos contemplados a los que corresponden. Podemos comprobar mediante la inspección de los datos que los grupos de aplicación han diferido tanto en el número como en las actividades que finalmente han puesto en práctica durante la fase experimental. Es destacable el bajo número de actividades que pudieron ser realizadas por el grupo A de aplicación. Ya comentamos en páginas anteriores que las condiciones en las que se desarrolló el trabajo con este grupo fueron especialmente difíciles, debido, por una parte, a factores relativos a las circunstancias propias del centro y su entorno, como el elevado absentismo de algunos de los componentes del grupo, así como, por otra, a las dificultades que tuvo en la práctica la profesora colaboradora para mantener una mínima continuidad en la aplicación del dispositivo, por la necesidad de atender otras necesidades del centro.

Otro aspecto destacable es que, como podemos ver, no se lograron realizar por parte de ninguno de los grupos de aplicación, las actividades pertenecientes a los dos últimos bloques de procesos contemplados en la teoría de base del dispositivo. Estos dos bloques eran *Procesos metacognitivos*, donde se encontraban las tres actividades específicas propuestas para la valoración de los metaconocimientos y los procesos de autorregulación, y *Procesos de ajuste personal-social*, donde se encontraba la actividad mediante la cual se evaluaba el contexto de aula. Lógicamente, una aplicación no regida por la dinámica impuesta por un diseño de investigación, a través del cual se pretendía ejercer un cierto control sobre determinadas variables, entre las cuales se hallaba la secuencia de actividades realizada, habría conllevado probablemente la realización pronta de las tareas específicas relacionadas con la metacognición. Así ocurrió en la aplicación piloto del dispositivo EDPL, realizada con base en el modelo de aplicación propuesto originalmente para el profesorado y los equipos de orientación educativa. En este modelo, obviamente, la preocupación por seguir una determinada secuencia de actividades o por realizar necesariamente un determinado grupo de ellas no existe, centrando el evaluador su esfuerzo en la obtención de información valiosa orientada a la intervención, a raíz de la aplicación de aquellas actividades del dispositivo que se consideran oportunas en cada momento en función de las hipótesis previas o de la información continua que van aportando a su vez cada una de las actividades realizadas.

Aunque estamos seguros de que la realización de todas las tareas del dispositivo, y especialmente de aquellas que específicamente evaluaban los procesos metacognitivos implicados en la comprensión, hubiera aportado datos relevantes a nuestro estudio, el hecho de que finalmente el tiempo con el que se contaba no nos haya permitido obtener esta información, no debía suponer un menoscabo significativo a las conclusiones que

Tabla 69. Actividades realizadas por cada aplicador. Las casillas coloreadas indican las tareas puestas en práctica.

Grupos de aplicación	Procesos psicológicos subyacentes									Procesos de A. G-F			Procesos de Integración textual										Procesos de Integración texto-conocimientos					Total	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27		28
Grupo A CEIP Andalucía																													6
Grupo B CEIP A. Machado																													15
Grupo C CEIP El Ruedo																													12
Grupo D CEIP El Ruedo																													12
Grupo E CEIP Lope de Rueda																													14
Grupo G CEIP Sánchez Alonso																													14
Grupo G IES L. de Guzmán																													9
Grupo H IES L. de Guzmán																													10
Grupo I IES Mariana Pineda																													9
Grupo J IES Torreblanca																													13
Grupo K IES Torreblanca																													10
<i>Totales</i>	6	1	1	1	1	1	0	3	3	11	2	4	11	3	0	11	10	10	9	6	3	8	7	4	3	2	1	2	124

La información recogida en esta tabla no recoge las actividades 29, 30 y 31, pertenecientes al bloque de Procesos metacognitivos, y 32, perteneciente al bloque de Procesos de ajuste personal-social (evaluación del contexto de aula). Estas actividades no pudieron ser puestas en práctica finalmente por ninguno de los aplicadores.

podían efectuarse. En efecto, dada la naturaleza metacognitiva del dispositivo EDPL, la evaluación de la puesta en práctica de los metaconocimientos y de los procesos de autorregulación, tiene lugar desde el inicio del proceso. Como hemos puesto de manifiesto durante el análisis del instrumento, todas las actividades propuestas contienen de forma preferente indicadores de evaluación relativos a los procesos de autorregulación o a los metaconocimientos sobre la lectura, y así mismo, los patrones metodológicos de evaluación dinámica propuestos, inciden especialmente en la observación y valoración cualitativa, por parte de los aplicadores, de la puesta en práctica de habilidades metacognitivas en el desarrollo de las tareas por parte del alumno. Por tanto, de cada una de las tareas realizadas ha podido extraerse abundante información sobre indicadores de evaluación de los procesos mencionados.

La cuantificación de las valoraciones en EDPL

La aplicación experimental del dispositivo de evaluación dinámica EDPL requería de los profesores-aplicadores la valoración, para cada una de las actividades realizadas, de una serie de indicadores que estaban recogidos al final de cada actividad, a modo de criterios de evaluación. Estos indicadores, que quedaron recogidos en el cuadro 30, sirven fundamentalmente al propósito de facilitar el análisis del proceso de resolución de la actividad propuesta y guardan relación con los procesos a evaluar de forma preferente en cada una de las tareas propuestas. La valoración de estos indicadores debía ser sobre todo cualitativa y orientada al establecimiento de posibles soluciones a las dificultades encontradas. Durante la fase experimental de nuestra investigación y al finalizar la misma, creímos oportuno desarrollar también un sistema de análisis que cuantificara de algún modo las valoraciones efectuadas por los aplicadores con relación a los indicadores. Esto nos permitiría operativizar el proceso de aplicación del dispositivo y disponer de unas puntuaciones para cada actividad y, así mismo, para la totalidad del dispositivo. De esta forma, la «puntuación» en EDPL se obtendría al operativizar el proceso de evaluación, mediante la asignación de una puntuación en función de la ejecución de las distintas actividades. En este sentido, en su reciente revisión sobre los procedimientos para el análisis de la información, Lukas y Santiago (2004) destacan la posibilidad demostrada de analizar cuantitativamente datos de naturaleza cualitativa. Tal y como apuntábamos anteriormente, hemos tratado de analizar los datos de nuestro estudio mediante la adopción de un enfoque flexible, que pudiera aportar una perspectiva de análisis cualitativo a los datos más cuantitativos, e igualmente, una perspectiva de análisis cuantitativo de datos en principio de naturaleza cualitativa.

El sistema adoptado para asignar un valor cuantificable a los datos extraídos de la valoración realizada por los aplicadores de EDPL, fue tomado de Mora (1998). En su investigación para evaluar el programa de enriquecimiento cognitivo *Comprender y Transformar*, el autor presenta una plantilla-registro ideada y utilizada originalmente para valorar las conductas cognitivas de los sujetos que participaban del programa. Esta valoración era efectuada por los aplicadores del citado programa tras el visionado de videos y con objeto de facilitar el análisis de los mismos. Las plantillas-registro contenían los indicadores de conductas cognitivas enriquecidas, agrupados por procesos o bloques, y a su vez una rejilla donde los evaluadores debían valorar el grado de consecución de cada uno de los indicadores en función de una escala que servía al propósito de calificar las conductas del grupo o alumno observado. Mediante este *registro de índices de conductas cognitivas enriquecidas* (Mora y Mora Merchán, 1995,

1998) puede efectuarse una valoración grupal o individual de los sujetos, utilizando una pequeña escala compuesta por cuatro apreciaciones (aunque originalmente los autores otorgan un valor 0 a la ausencia de conducta, nosotros hemos optado por asignar a esta condición valor 1. Las razones no van más allá de facilitar el manejo de los datos):

- 1 No aparece la conducta
- 2 Indicios. Conducta rudimentaria
- 3 Conducta claramente apreciable (aunque no sea de gran calidad)
- 4 Conducta intensa o de gran calidad

Esta valoración cualitativa del desempeño de los sujetos en el programa Comprender y Transformar, no se traduce en su propósito original en unas puntuaciones con las que establecer análisis estadísticos, o con las que reinterpretar los datos de las observaciones. La asignación de un valor numérico no tenía más significado que el de facilitar la codificación de los aplicadores en función de la adecuación de la conducta a uno de los valores establecidos. Sin embargo, dado el claro carácter gradual de los cuatro valores, la posibilidad de asignar un valor numérico real que se correspondiera con cada uno de ellos, podría facilitar un doble análisis de nuestros datos; por una parte, el análisis cualitativo que se desprende de la valoración de los profesores aplicadores sobre la puesta en práctica de cada uno de los indicadores propuestos para evaluar la actividad; por otra, el análisis cuantitativo que se desprende de la asignación de valores numéricos a estas valoraciones efectuadas por los aplicadores sobre los citados indicadores. Como señalábamos en líneas anteriores y reseñan entre otros autores Lizasoain y Joaristi (2000), citados en Lukas y Santiago (2004): *La evaluación (...) se caracteriza por la complementariedad de forma que el análisis de datos debe llevarse a cabo utilizando de forma complementaria o simultánea múltiples enfoques, métodos y técnicas.*

Una vez descrito de forma sucinta el sistema que adoptamos para asignar un valor cuantitativo a las observaciones y valoraciones efectuadas por los aplicadores durante el desarrollo de las sesiones del dispositivo, nos disponemos ahora a exponer el procedimiento que hemos seguido para la cuantificación de los datos y para el posterior análisis.

En primer lugar se procedió a insertar en cada una de las *Hojas de registro y evaluación* confeccionadas para cada una de las actividades propuestas en el dispositivo, una rejilla que serviría al propósito de puntuar cada uno de los indicadores contemplados en la actividad en función de las observaciones y valoraciones efectuadas por el aplicador. De este modo, esta «nueva» Hoja de registro (Anexo XIV) incorpora una de las sugerencias realizadas por los aplicadores en su valoración del dispositivo, y ha sido la base sobre la que se ha realizado el análisis cuantitativo del proceso de aplicación de EDPL. Posteriormente, se han revisado minuciosamente cada uno de los 60 *Cuadernos de hojas de registro y evaluación* entregados a los profesores aplicadores. Estos cuadernos contenían todas las hojas de registro de las actividades propuestas, y en éstas los aplicadores han volcado la información procedente de su análisis cualitativo, efectuado sobre cuatro aspectos fundamentalmente: (a) la evaluación del proceso de resolución de la actividad, (b) la mediación efectuada, (c) la evaluación de procesos metacognitivos, y (d) las observaciones que consideraran oportunas. El sistema utilizado para la asignación de los distintos valores cuantitativos considerados, ha seguido el siguiente patrón:

* Se ha otorgado 1 punto (*No aparece la conducta*) en las siguientes circunstancias: (a) cuando el aplicador hacía referencia explícita a la ausencia de la puesta en práctica del indicador. Ejemplos de estas valoraciones extraídos de los protocolos de evaluación serían: *no controla la impulsividad; no persiste cuando se enfrenta a obstáculos o dificultades*, etc.; (b) cuando el aplicador no hace referencia al indicador.

* Se han otorgado 2 puntos (*Indicios de conducta o puesta en práctica del indicador de forma rudimentaria*) cuando el aplicador refiere explícitamente la puesta en práctica del indicador por parte del sujeto, mencionando sin embargo un nivel de ejecución de baja calidad o incipiente, o haciendo referencia a la necesidad de ayuda para la consecución del indicador. Ejemplos de esta valoración podrían ser: *necesita ayuda para evaluar logros con relación a las metas definidas; se distrae con facilidad, está bastante inquieto; pide que repita algunas palabras...*

* Se han otorgado 3 puntos (*Conducta claramente apreciable –aunque no sea de gran calidad–*) cuando de forma explícita el aplicador menciona la puesta en práctica del indicador. Ejemplos de esta valoración son: *realiza adecuadamente el seguimiento visual de un texto sin saltarse palabras ni líneas; almacena, recupera y utiliza información de manera simultánea a la realización de la tarea; identifica la tarea que debe realizar y el problema planteado; etc.*

* Se han otorgado 4 puntos (*Conducta intensa o de gran calidad*) en las siguientes circunstancias: (a) cuando la observación o valoración del aplicador hace referencia explícita a la alta calidad de la conducta manifestada o a su intensidad. Como ejemplos de esta valoración podemos reseñar valoraciones del tipo: *Utiliza la mediación verbal en todo momento; controla la actividad durante la lectura para determinar si entiende o no lo que está leyendo, buscando activamente la transformación de las pseudopalabras en palabras con significado; la puesta en práctica de habilidades de supervisión y autoevaluación por parte de este alumno ha influido notablemente en que sus resultados en esta tarea –que representa muchas dificultades para él– sean significativamente mejores; conoce la estructura fonológica de las palabras... muy bien en las transformaciones de palabras;* (b) cuando el aplicador refiere de forma explícita la puesta en práctica del indicador, transcribiendo literalmente el contenido propuesto en el Manual de trabajo, y la ejecución manifiesta del alumno, recogida en las hojas de registro y evaluación o en anotaciones aparte, puede ser considerada de gran calidad.

A pesar de que, como hemos ya apuntado cuando abordamos el procedimiento seguido en la investigación, se incidió especialmente durante las sesiones del Grupo de trabajo en la conveniencia de que los aplicadores valoraran el proceso de aplicación de las actividades del dispositivo haciendo uso de los indicadores propuestos para cada tarea, esto no siempre pudo ser así. En este sentido, algunos profesores aplicadores no recogieron en las *Hojas de registro y evaluación* la valoración relativa a la puesta en práctica de la totalidad de los indicadores, haciendo alusión a la mayoría de éstos pero dejando de valorar otros. Por otra parte y de forma general, la gran mayoría del material registrado por los aplicadores, con base en el cual se ha efectuado gran parte del análisis cualitativo que exponemos, sí ha recogido las valoraciones sobre el proceso de mediación de cada actividad realizada, haciendo hincapié también en la evaluación de los procesos metacognitivos implicados.

El hecho de que no todos los indicadores de evaluación hubieran sido valorados por los aplicadores, introducía un problema en la consecución del sistema de codificación explicitado anteriormente. Era necesario conocer las razones por las cuales se dejaron de valorar algunos de los indicadores propuestos, con el fin de encontrar una vía de solución que nos permitiera aplicar el sistema minimizando los problemas. Los aplicadores adujeron las siguientes razones: (a) se trataba de indicadores que no fueron evaluados en esa actividad; (b) fueron considerados menos relevantes a la hora de observar y valorar la actividad; (c) se trató de indicadores no observados durante el proceso de evaluación; (d) se trataba de indicadores ya evaluados y valorados en actividades anteriores, por lo que no se repitió la valoración.

Fueron varias las posibilidades que se barajaron para «completar» la valoración de cada una de las actividades en las que había indicadores que no habían sido valorados. Una de estas posibilidades suponía dejar la casilla de la rejilla en blanco y no contabilizar el indicador en las medias resultantes. Esta opción respondía sobre todo a la opinión de los aplicadores que hacía referencia a que los indicadores para los cuales no se había emitido una observación no habían sido evaluados. Por otra parte, otras dos opciones contemplaban la codificación de los indicadores no valorados, bien con 2,5 puntos, al ser ésta la media de las puntuaciones posibles, o bien con la media de la puntuación de esos mismos indicadores a lo largo del dispositivo. Estas dos opciones respondían mejor a las opiniones que indicaban que los indicadores habían sido ya evaluados en actividades anteriores. Por último, una tercera opción era otorgar la puntuación mínima a los indicadores no valorados. Esta opción, finalmente elegida, respondía a las opiniones que expresaban que la ausencia de valoración fue debida a que los indicadores no fueron observados durante el proceso de mediación o, sencillamente, no fueron evaluados. Las razones que nos llevaron finalmente a optar por este criterio, fueron primordialmente dos: por una parte, otorgar la puntuación mínima posible ante la ausencia de valoración por parte del aplicador, garantizaba que no se inflaran artificialmente las puntuaciones obtenidas por los sujetos en su actividad; por otra, la puntuación mínima parecía ser justa por cuanto el aplicador no apreció la puesta en práctica manifiesta de la conducta durante el desarrollo de la actividad, por lo que habría que pensar que, en la mayoría de los casos, la conducta no apareció. De esta forma, durante el proceso de codificación y cuantificación de los protocolos de evaluación dinámica, asignamos 1 punto a los indicadores que no fueron valorados expresamente por los aplicadores.

Una vez realizada la operación anterior, se procedió a tipificar las puntuaciones, construyendo con éstas una escala derivada similar a la elaborada para las pruebas de comprensión y actitudes, es decir, con puntuación media en 100 y desviación típica en 16. Las razones que nos llevaron a la transformación de las puntuaciones directas obtenidas de la valoración cuantitativa del dispositivo EDPL, fueron similares a las ya expresadas para llevar a cabo en su momento la transformación en las pruebas criterio, teniendo en este caso un mayor peso la idoneidad de adoptar un marco estándar de referencia mediante el que comparar las puntuaciones obtenidas por distintos sujetos en distintos grupos de aplicación. Se conseguía de esta forma reducir el impacto que podría tener en los contrastes y en los resultados, la diferente medida valorativa efectuada por cada aplicador en las actividades. Esta opción adoptada en el análisis, unida al hecho de que la cuantificación de los datos cualitativos registrados por los aplicadores, fue realizada por un solo sujeto, minimizando o eliminando de esta forma la variabilidad inter-sujeto en esta valoración, otorgaba mayores garantías al proceso de análisis.

Otro aspecto relevante que surgió durante el desarrollo del análisis de los datos contenidos en las *Hojas de registro y evaluación* del dispositivo, fue el hecho de que algunos aplicadores realizaron valoraciones sobre otros indicadores, algunos de ellos propuestos en el Manual, pero no recogidos explícitamente en la actividad, y otros no contemplados en la propuesta recogida en el Manual. En la tabla que exponemos a continuación se recoge la totalidad de indicadores propuestos en cada una de las actividades que finalmente fueron realizadas. Así mismo, posteriormente hacemos referencia al número de indicadores pertenecientes a procesos valorados por los aplicadores aunque no estuvieran inicialmente recogidos en la actividad que estaban evaluando.

Tabla 70. Número de indicadores de evaluación propuestos en cada uno de los procesos para cada una de las actividades del dispositivo EDPL

Act.	Procesos Psicológicos Subyacentes	Asociación Grafema-Fonema	Integración textual	Integración texto-conocimientos	Procesos Meta-cognitivos	Ajuste personal-social	TOTAL
1	3		1		9	2	15
2	2				10		12
3	3				6		9
4	3				6		9
5	3				11		14
6	3				6		9
8	3			1	6	1	11
9	3				8		11
10	3	2			4	2	11
11	2	3			6	1	12
12	1	3			8	4	16
13	2	2	1		4	2	11
14	1		1		4		6
16	2		1		6	2	11
17	2		3		6	1	12
18	1		2		8	3	14
19			1		9	1	11
20	1		3		6	2	12
21	2		3		5	1	11
22	3		4	2	9	1	19
23			9	9	15	3	36
24			4	7	11	3	25
25	1		1	3	8	1	14
26	1		1	2	6		10
27				7	14		21
28			1	8	12		21
Total	45	10	36	39	203	30	363

A pesar de que el número total de veces que fueron utilizados los indicadores de evaluación no propuestos inicialmente en las actividades fue de 32, el número de sujetos que fueron valorados con estos indicadores fue sólo de 22. Esto responde a que hasta en seis ocasiones un alumno fue valorado con más de un indicador no propuesto inicialmente. Podemos observar que los aplicadores necesitaron utilizar más indicadores en tres de los procesos: Procesos psicológicos subyacentes, Integración texto-conocimientos, y Ajuste personal-social; mientras que no fue así en los Procesos de

Tabla 71. Número de veces que fueron utilizados los indicadores de evaluación pertenecientes a procesos que no figuraban inicialmente en las actividades

Act.	Procesos Psicológicos Subyacentes	Asociación Grafema-Fonema	Integración textual	Integración texto-conocimientos	Procesos Meta-cognitivos	Ajuste personal-social	TOTAL
1				2			2
2						5	5
6						1	1
9						1	1
13				1			1
14						3	3
19	4			2			6
23	6						6
24	2						2
26						2	2
28	1					2	3
Total	13	0	0	5	0	14	32

asociación G-F, los procesos de Integración textual, y los Procesos metacognitivos. De esta forma, los profesores aplicadores habrían contribuido a completar el listado de indicadores de evaluación propuestos para cada una de las actividades. Es necesario decir, sin embargo, que en una propuesta de evaluación dinámica como la nuestra, este listado de indicadores propuestos siempre tendría necesariamente un carácter «incompleto», ya que en función del contexto de aplicación (evaluador, dificultades que se detecten, respuesta del sujeto a la mediación, etc.) el aplicador puede considerar oportuno la inclusión de indicadores *extra* para valorar de una forma más completa la situación planteada. Fundamentalmente, el contenido de los indicadores de los procesos psicológicos subyacentes recogidos en la tabla anterior, ha estado relacionado con el mantenimiento de la atención durante el desarrollo de las actividades, criterio que fue recogido en la mayoría de las tareas, pero que ha sido valorado también por los aplicadores en las que no figuraba. En cuanto al bloque de procesos de integración texto-conocimientos, los aplicadores han necesitado utilizar indicadores del tipo: *justifica respuestas y decisiones*, que no se encontraban inicialmente en todas las actividades. Por último, los indicadores relativos al bloque de procesos de ajuste personal-social han estado relacionados con la actitud, la motivación y el interés mostrado durante el desarrollo de las actividades de lectura en general, y de las tareas propuestas en particular. Los aplicadores mostraron un gran interés en reflejar en sus valoraciones, tanto aquellas expresadas en las *Hojas de registro y evaluación*, como las expuestas en la evaluación conjunta de la aplicación del dispositivo realizada en el Grupo de Trabajo, estos aspectos de tipo socio-emocional, considerándolos *clave* en la respuesta a la acción educativa ejercida en la aplicación de EDPL.

Los anteriores indicadores fueron valorados por los aplicadores en los mismos términos que los propuestos inicialmente, siendo considerados por nosotros de gran interés para el proceso de evaluación. Como quedó recogido cuando se analizó el contenido del dispositivo, en cada actividad aparecían una serie de indicadores de evaluación que conectaban especialmente con los procesos a evaluar en la actividad o que eran considerados relevantes en el proceso de valoración. Al final de la lista de indicadores en cada actividad se recogía un párrafo que aludía a la posibilidad de utilizar otros indicadores además de los contemplados en el dispositivo y que se adecuaran al proceso y al contexto de evaluación seguido. Esta posibilidad contribuía a contextualizar nuestra

propuesta de evaluación, dejando abierta la posibilidad de utilizar criterios considerados oportunos por el aplicador en función de cada caso concreto. La cuestión problema que se nos planteaba en el estudio era cómo tenerlos en cuenta en el análisis. La eliminación sin más de estas valoraciones en el análisis hubiera sido incoherente con un proceso de evaluación dinámica que pretende primordialmente obtener información relevante de cara a la acción educativa eficaz; por tanto, la apreciación de determinados procesos o habilidades por parte del aplicador debían ser tomados en consideración como parte importante de su análisis. De esta forma, cada actividad en función del aplicador, ha podido tener un diferente número de indicadores de evaluación valorados y posteriormente cuantificados. Finalmente, los análisis han sido efectuados para cada actividad, tomando en consideración la totalidad de los indicadores propuestos y también aquellos que inicialmente no formaban parte de la actividad pero fueron valorados por los aplicadores. Este análisis se ha realizado en función de los distintos procesos contemplados en el dispositivo, agrupados a su vez en las tres categorías ya citadas con anterioridad y que ahora recordamos:

(a) *Procesos metacognitivos*: Procesos de conocimiento consciente sobre la lectura, sobre las estrategias que pueden usarse, y sobre procesos de ajuste personal-social (metaconocimientos); y Procesos de autorregulación de la comprensión;

(b) *Análisis e integración de la información*: Procesos implicados en la asociación de grafemas y fonemas; Procesos psicológicos subyacentes; Procesos implicados en la integración textual; y Procesos de integración texto-conocimientos previos; y

(c) *Procesos implicados en el Ajuste Personal-Social*.

De este modo, se han calculado y considerado diversas puntuaciones en función de distintos criterios: (a) puntuación obtenida en la actividad, calculada a raíz de las valoraciones cuantitativas efectuadas sobre cada indicador; (b) puntuación media obtenida en los indicadores que especialmente se relacionaban con los procesos a evaluar contemplados en cada actividad; (c) puntuación, para cada actividad, en cada uno de los 7 procesos contemplados en el dispositivo, obtenida a raíz del promedio de los valores asignados a los indicadores pertenecientes a los distintos procesos; (d) puntuación global obtenida en cada uno de los procesos contemplados; (e) puntuación global obtenida en EDPL, obtenida a raíz del promedio de las puntuaciones en cada uno de los procesos contemplados; y, por último, (f) puntuación obtenida en función de otros criterios específicos para un grupo de actividades.

La matriz de datos en EDPL

La matriz de datos construida para el análisis cuantitativo y cualitativo de la evaluación dinámica, a raíz de las valoraciones efectuadas por los aplicadores en las *Hojas de registro y evaluación*, contenía por consiguiente la totalidad de los indicadores propuestos inicialmente, así como aquellos otros que fueron utilizados por los evaluadores durante el desarrollo de las actividades y que no se encontraban inicialmente en la propuesta. Como hemos indicado anteriormente, los indicadores eran puntuados de 1 a 4 puntos en función de las observaciones y valoraciones efectuadas por los aplicadores. Así mismo, la matriz contenía para cada actividad, los procesos contemplados en el dispositivo que eran a su vez evaluados en la tarea. De esta forma se podía acceder al análisis de la posible evolución de un sujeto o un grupo de sujetos en

un grupo determinado de indicadores o en un grupo de procesos. La matriz contenía también variables específicas para algunas tareas, que respondían al interés que podría tener una valoración más cuantitativa del contenido de las mismas. De este modo, algunas actividades que, además de la valoración cualitativa de la puesta en práctica de los indicadores, podían ser valoradas en términos de aciertos o errores en los ítems propuestos, contenían variables que recogían las distintas puntuaciones que podían obtenerse en su ejecución. Para algunas de estas actividades también se recogía una puntuación pretest, obtenida inicialmente y previa al proceso de mediación del aplicador, y una posttest, obtenida con posterioridad al proceso de mediación.

Por otra parte, los aplicadores, en sus observaciones y valoraciones efectuadas sobre el proceso de aplicación, hacían alusión a menudo a la necesidad que un sujeto tenía de la mediación para poder ejecutar mejor una determinada tarea. Así mismo hacían referencia a la calidad de la respuesta del sujeto a la mediación ofrecida, de modo que algunos alumnos parecían beneficiarse claramente de la mediación del evaluador, pudiendo obtener puntuaciones más altas en los mismos indicadores de posteriores actividades. Sin embargo, estas referencias a la mediación que un sujeto requería para poder realizar con éxito una determinada tarea, podían ofrecer también una medida indirecta de las dificultades presentes en ese mismo sujeto. De esta forma, el número de veces que un sujeto se beneficiaba de la mediación del aplicador podría ser un indicador de la presencia de mayores dificultades y, por tanto, ser proporcional a valoraciones y puntuaciones menores en el dispositivo. La posibilidad de realizar este análisis quedó reflejada en las *Hojas de registro y evaluación* mediante una cruz que acompañaba a la puntuación asignada y, así mismo, quedó reflejada posteriormente en la matriz mediante una variable que aludía al número de veces que el sujeto se benefició de la mediación del evaluador. La hipótesis de trabajo relacionada con esta variable apuntaba por tanto en una doble dirección con sentido contrapuesto. Por una parte, podía pensarse que aquellos sujetos que se hubieran beneficiado de la mediación y que, por tanto, hubieran obtenido mejores puntuaciones en esta variable, verían incrementadas sus puntuaciones finales en el dispositivo; por otra, también era lógico esperar que los sujetos que más hubieran necesitado de la intervención del aplicador durante el desarrollo de las tareas, serían aquellos que más dificultades presentarían y, a la postre, los que obtendrían valoraciones más pobres y menores puntuaciones en el dispositivo.

Se han analizado 60 *Cuadernos de Hojas de Registro y Evaluación*, pertenecientes a los 60 sujetos que han participado finalmente de la fase experimental en el grupo de tratamiento. Si observamos la tabla recogida anteriormente, el promedio de actividades realizadas por cada sujeto sería de unas once, lo que vendría a sumar más de 650; sin embargo, el número total de actividades que han sido revisadas y valoradas finalmente, ha sido de 598. No podemos olvidar que algunos sujetos tuvieron un índice de absentismo alto, por lo que, en ocasiones, aunque el grupo de aplicación realizara en su conjunto un número determinado de actividades, algunos alumnos sólo pudieron ser evaluados de algunas de ellas. Por otra parte, tal y como expresábamos con anterioridad, en la mayoría de los grupos se realizó una aplicación colectiva de las tareas propuestas en el dispositivo, y en algunos, un análisis también conjunto de todas o algunas de las actividades realizadas. A pesar de esto, la valoración cualitativa realizada por los aplicadores ha sido llevada a efecto sobre cada uno de los sujetos participantes, lo que ha facilitado la cuantificación de estas observaciones.

Los resultados obtenidos en el dispositivo

En primer lugar nos disponemos a mostrar los resultados globales para el GE (N = 60) de cada uno de los procesos contemplados en el dispositivo, así como la puntuación dinámica global. La transformación de las puntuaciones en la forma a la que hicimos alusión anteriormente, ocasiona lógicamente que los promedios obtenidos para cada proceso tengan un mismo valor (100) y, así mismo, ocurre con los valores de las desviaciones típicas (16). En este caso, con el fin de poder valorar los puntajes medios diferenciales obtenidos en cada proceso y en el total, hemos anotado los promedios obtenidos previamente a la transformación realizada. Estos valores promedio están calculados lógicamente sobre la puntuación máxima que podía obtenerse, es decir, 4 puntos. De esta forma, tratándose de puntuaciones equivalentes a las transformadas, podremos entender mejor algunos datos correlacionales que aparecerán posteriormente. En la tabla siguiente se recogen los valores de correlación obtenidos entre los procesos y con la puntuación global; el análisis de estas correlaciones puede ayudarnos a determinar, por una parte, la coherencia interna del contenido del dispositivo EDPL, ya que en nuestra teoría de base propugnamos un modelo interactivo y sistémico de lectura, en el que los distintos procesos que intervienen en su aprendizaje y dominio se encuentran interrelacionados. Por otra parte, también debe ayudarnos a determinar la consistencia y fiabilidad del procedimiento de valoración efectuado.

Tabla 72. Valores promedio de las puntuaciones dinámicas en cada uno de los procesos contemplados en EDPL para el GE en primaria (N = 32) y en secundaria (N = 28), y Valor promedio global

Procesos contemplados en el dispositivo EDPL	Valor promedio Primaria	DT	Valor promedio Secundaria	DT	Valor promedio global	DT
Procesos Psicológicos Subyacentes	2,738	0,335	3,070	0,222	2,893	0,330
Procesos de Asociación G-F	3,188	0,463	3,402	0,362	3,288	0,429
Procesos de Integración Textual	2,579	0,390	2,918	0,347	2,737	0,405
Procesos de Integración Texto-Conoc.	1,861	0,711	1,427	0,733	1,658	0,748
Procesos Metacognitivos	2,177	0,395	2,381	0,372	2,272	0,395
Procesos de Ajuste Personal-Social	2,074	0,571	2,336	0,475	2,197	0,540
Puntuación Dinámica Global	2,328	0,372	2,575	0,324	2,443	0,369

En el gráfico que mostramos a continuación podemos observar estas diferencias de forma más nítida. En efecto, las tres primeras barras, pertenecientes a los Procesos psicológicos subyacentes, procesos de asociación G-F y procesos de integración textual, obtienen diferencias importantes en cuanto a los promedios alcanzados con relación a las otras tres: los procesos de integración texto-conocimientos previos, procesos metacognitivos y procesos de ajuste personal-social.

Figura 5. Valores promedio para cada uno de los procesos contemplados en EDPL

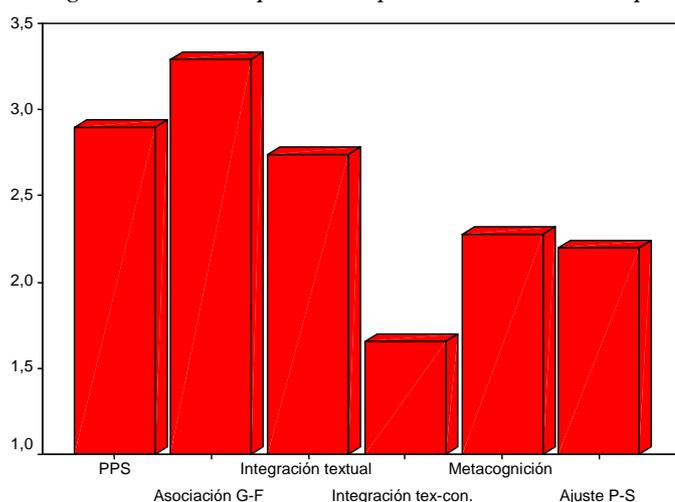


Tabla 73. Matriz de correlaciones entre las puntuaciones obtenidas en los distintos procesos implicados para el GE completo (N = 60)

	P.P.S.	Asociación G-F	Integración textual	Integración texto-con.	P. Meta-cognitivos	Ajuste P-S	Puntuación dinámica G
P.P.S.		0,364**	0,643**	0,078	0,710**	0,659**	0,812**
Asociación G-F	0,364**		0,496**	-0,201	0,159	0,094	0,290*
Integración textual	0,643**	0,496**		0,130	0,593**	0,314*	0,671**
Integración texto-con.	0,078	-0,201	0,130		0,425**	0,375**	0,371**
P. Meta-cognitivos	0,710**	0,159	0,593**	0,425**		0,766**	0,972**
Ajuste P-S	0,659**	0,094	0,314*	0,375**	0,766**		0,818**
Puntuación dinámica G	0,812**	0,290*	0,671**	0,371**	0,972**	0,818**	

La inspección de las puntuaciones medias absolutas mostradas en la primera de las tablas nos puede ayudar a determinar qué procesos son los que han representado una mayor dificultad para los sujetos. En este sentido y teniendo en cuenta que el valor medio posible es de 2,5 puntos, tres procesos (*integración texto-conocimientos previos*, *procesos de ajuste personal-social* y *procesos metacognitivos*) son los que se sitúan por debajo de este nivel; tanto si analizamos los resultados en función de la etapa como si lo hacemos de una forma global. El primero de ellos ha obtenido un promedio que se encuentra significativamente por debajo del resto, mostrando igualmente una mayor variabilidad en sus puntuaciones. Estos datos se ven a su vez reflejados en la matriz de correlaciones presentada, obteniendo en general valores de correlación bajos con relación al resto de procesos e incluso negativo con uno de ellos. Paradójicamente, en la muestra perteneciente a educación secundaria, los *procesos de integración texto-conocimientos previos*, a pesar de presentar un valor absoluto muy bajo, obtienen correlaciones más altas con el resto de puntuaciones. Por otra parte, los *procesos de ajuste*, como recordaremos, hacen referencia, entre otros, a aspectos relativos a la

motivación, las actitudes, o el interés mostrado en las actividades de lectura; muchos profesores consideran que estos aspectos constituyen uno de los problemas clave en la explicación de las dificultades relacionadas con la lectura, especialmente cuando se trata de población en situación de desventaja socio-cultural. Nuestros datos parecen haber reflejado esta realidad, confirmada a la sazón en las sesiones del Grupo de Trabajo por los aplicadores del dispositivo. En línea con lo anterior, podemos decir que gran parte de los indicadores que han sido utilizados por los aplicadores sin que hubieran sido recogidos inicialmente para valorar la actividad, pertenecieron al bloque de procesos de ajuste personal-social; los aplicadores destacaron, tanto en las sesiones de trabajo conjunto como en sus *Hojas de registro y evaluación*, la relevancia de estos procesos en la explicación de los problemas manifiestos en lectura y comprensión. El segundo bloque de procesos al que los aplicadores acudieron mayoritariamente para extraer indicadores que inicialmente no estaban recogidos en las actividades (en este caso, hacemos referencia a indicadores propuestos en el Manual y pertenecientes a procesos que eran evaluados en la tarea, aunque inicialmente no estuvieran recogidos en la actividad), fue el de procesos metacognitivos. También para este bloque los resultados muestran la existencia de dificultades, dato que confirmaron de forma continuada los aplicadores en las sesiones de trabajo. Es destacable, sin embargo, que para los bloques de ajuste y de procesos metacognitivos, especialmente para este último, los valores de las correlaciones establecidas con el resto de procesos sean significativos o muy significativos y con valores medios y altos, exceptuando a los *procesos de asociación G-F* con quien mantienen una correlación positiva pero baja. También son estos dos procesos los que mantienen una correlación más alta y significativa con la puntuación dinámica global. Estos datos apuntan a la relevancia de estos procesos en la configuración final del proceso de lectura y comprensión, así como, a tenor de sus puntuaciones, en la configuración de las principales dificultades de aprendizaje en la lectura. Este dato, y muy especialmente el que hace referencia a los procesos metacognitivos, vendría a apoyar nuestra quinta hipótesis, relativa a la relevancia de la metacognición en la explicación de las dificultades de aprendizaje de la lectura.

Es destacable así mismo el alto número de correlaciones significativas y muy significativas entre los distintos procesos y entre éstos y la puntuación dinámica global. Como apuntábamos anteriormente, este dato refuerza notablemente la consistencia del dispositivo en cuanto a su contenido y a la teoría de base que lo sustenta, así como también el sistema de codificación y valoración utilizado. De esta forma, los procesos contemplados pertenecen a un mismo conjunto de elementos que contribuyen a su vez a valorar un determinado constructo, en este caso la comprensión. Incluso aquellos procesos que podrían considerarse a priori menos relacionados directamente con la comprensión, como los *procesos psicológicos subyacentes*, los *procesos de asociación G-F* o los *procesos de ajuste personal-social*, muestran unos niveles de correlación aceptables, en el caso de la asociación G-F, o incluso muy altos en el caso de los otros dos procesos. En este sentido, los distintos procesos, mostrando los aspectos diferenciales que analizan, contribuyen de forma significativa a la evaluación global de la lectura y la comprensión.

Hemos efectuado un análisis de regresión por pasos para evaluar la contribución de cada uno de los procesos a la puntuación dinámica global obtenida. Mediante el análisis obtuvimos una ecuación en la que figuraban 5 de los factores considerados, produciendo todos ellos cambios significativos en el valor R^2 . En primer lugar fue introducida la puntuación dinámica de los *procesos metacognitivos* con un valor R^2 de 0,944, una F de

982,895 y una probabilidad asociada de 0,000; en segundo lugar se incluye la puntuación dinámica de los *procesos psicológicos subyacentes*, incrementado el valor R^2 en 0,030; seguidamente se incluye la puntuación dinámica de los *procesos de asociación G-F*, produciéndose un incremento en R^2 de 0,006; en cuarto lugar la puntuación de los *procesos de ajuste personal-social*, con un incremento de 0,007; y finalmente se introdujo en la ecuación la puntuación de los *procesos de integración textual*, produciéndose un incremento en el valor R^2 de 0,004. El valor final de R^2 fue de 0,996; obviamente el modelo explicaba de forma muy significativa el resultado global. La puntuación de los *procesos de integración texto-conocimientos*, a pesar de mantener una correlación significativa con la puntuación dinámica global, no fue introducida en la ecuación; cuando procedimos a valorar la porción de varianza explicada de forma individual, obtuvimos que el valor R^2 era de 0,137, mostrando el contraste un valor significativo, con una F de 9,246 y una $p < 0,004$. Este método de introducción individual de los factores reveló que, además de la ya mencionada, todas las puntuaciones dinámicas contempladas explicaban significativamente el resultado global. La puntuación en los *procesos psicológicos subyacentes* obtenía un valor R^2 de 0,659, con una probabilidad asociada al valor F de $p < 0,000$; la puntuación en los *procesos de asociación G-F* explicaba de forma individual el 8,4%, siendo la que mostró un menor valor R^2 , aunque con una probabilidad asociada al valor F de $p < 0,024$; los *procesos de integración textual* obtenían una R^2 de 0,450, explicando de forma muy significativa el contraste con un valor de probabilidad $p < 0,000$; por su parte, los *procesos metacognitivos* fueron los que obtuvieron un valor R^2 mayor (0,944), explicando de forma muy significativa el contraste; por último, la puntuación obtenida en los *procesos de ajuste* obtuvo un valor R^2 de 0,668, también mostrándose capaz de explicar de forma muy significativa el contraste.

Por otra parte, en cuanto al nivel de puntuación alcanzado por los distintos procesos, uno de los que consigue un promedio que se sitúa por encima de la media es el de *procesos psicológicos subyacentes*. Como recordaremos, estos procesos hacían referencia a aspectos como la atención, la orientación espacial y temporal o la memoria. En general, y tomando el grupo experimental completo, estos buenos resultados, unidos al hecho de que la mayoría de los aplicadores no optó por realizar actividades de este bloque en respuesta a posibles dificultades relacionadas con los procesos mencionados anteriormente, nos hacen pensar que no se han encontrado aquí las mayores dificultades relacionadas con la lectura. Este dato fue confirmado por el profesorado aplicador en el transcurso de las sesiones del Grupo de Trabajo. Sin embargo, el bloque de procesos mejor puntuados, es el que hace referencia a los *procesos de asociación G-F*; también en este caso ha sido una minoría de los aplicadores los que han utilizado la posibilidad de realizar otras actividades de las propuestas en el bloque además de la que fue considerada «obligatoria» en el diseño de investigación. En este caso, especialmente para los aplicadores del dispositivo en secundaria, también se confirmaría el hecho de que estos procesos no estarían detrás de la mayor parte de las dificultades de lectura. Al igual que para los *procesos psicológicos subyacentes*, en las sesiones del Grupo de Trabajo, los aplicadores confirmaron mayoritariamente la ausencia de problemas significativos en estas habilidades. En efecto, podemos observar en la primera de las tablas expuestas, que para los sujetos de la muestra en secundaria, la puntuación más alta es precisamente la obtenida en los procesos de asociación G-F. Es destacable así mismo la falta de relación entre las puntuaciones de este bloque y el resto de los procesos, mostrando tres valores de correlación negativos y otros con valores muy bajos. Por otra parte, la correlación con la puntuación dinámica global en esta etapa es

de $r=0,013$, mientras que el resto de valores correlacionales para la puntuación global son muy significativos y con valores muy elevados.

Por último, este repaso al valor de las puntuaciones medias obtenidas por los distintos procesos, nos lleva ahora al bloque de *procesos de Integración textual*. Un aspecto que creemos relevante es que la puntuación promedio obtenida en el grupo completo por este proceso supera de forma significativa a la que muestra el bloque de procesos de *Integración texto-conocimientos previos*. Este dato o, mejor expresado, lo que puede deducirse de él, fue confirmado continuamente por los aplicadores, quienes destacaban las dificultades que suponía para los sujetos la puesta en práctica de las habilidades relacionadas con la realización de inferencias, la formulación de hipótesis sobre el contenido de los textos, el establecimiento de relaciones entre el contenido de los textos y los conocimientos previos, etc. Estas dificultades, aún existiendo, eran mucho menores cuando se trataba de organizar la estructura sintáctica de una frase, de relacionar imágenes con frases teniendo en cuenta la estructura sintáctica de las mismas, o incluso de utilizar la estructura de un texto para acceder a la comprensión del mismo.

En definitiva, con base en los resultados obtenidos en la evaluación dinámica de los procesos implicados en la lectura, podríamos decir a nivel general que dos son los procesos que representan una mayor dificultad para el alumnado que presenta dificultades de aprendizaje relacionadas con la lectura: por una parte, los procesos *implicados en la integración del texto con los conocimientos y experiencias previas*; por otra, los *procesos metacognitivos*. Así mismo, aunque no se trate de procesos relacionados directamente con aspectos lingüísticos, los resultados muestran la relevancia de los *procesos de ajuste personal-social* en la posible explicación de las dificultades de aprendizaje relacionadas con la lectura. Estos resultados se dan igualmente, aunque con ligeras diferencias, en las muestras de sujetos de primaria y de secundaria.

Hemos querido de alguna forma buscar confirmación de esto en las pruebas criterio de comprensión. Como se recordará, éstas estaban configuradas en distintas dimensiones que trataban de ofrecer no sólo un resultado global, sino una información lo más completa posible sobre los distintos procesos implicados en la comprensión. Hemos procedido para cada etapa al análisis de las correlaciones entre las distintas dimensiones contempladas en las pruebas y las puntuaciones dinámicas de los distintos procesos considerados en el dispositivo EDPL. Buscamos de esta forma valorar si las dificultades encontradas mayormente por los sujetos experimentales en el desarrollo de las actividades del dispositivo se ven confirmadas en las puntuaciones obtenidas en las pruebas estáticas de comprensión. Hemos establecido estas correlaciones con las puntuaciones pretest y posttest de las pruebas, con el fin de valorar también la posible evolución en la correlación observada. Lógicamente, tratándose del GE, esperábamos menores niveles de correlación entre las puntuaciones dinámicas y el pretest, aunque como tendremos oportunidad de comprobar y analizar más adelante, cuando profundicemos en el valor predictivo del dispositivo con relación a las distintas medidas observadas, las puntuaciones dinámicas, como valoraciones de un proceso de aprendizaje que son, podrían llegar a mostrar un nivel de independencia notable con relación a cualquier tipo de puntuación de carácter «estático», incluida la puntuación posttest. Destacamos a continuación los resultados más relevantes de estos análisis.

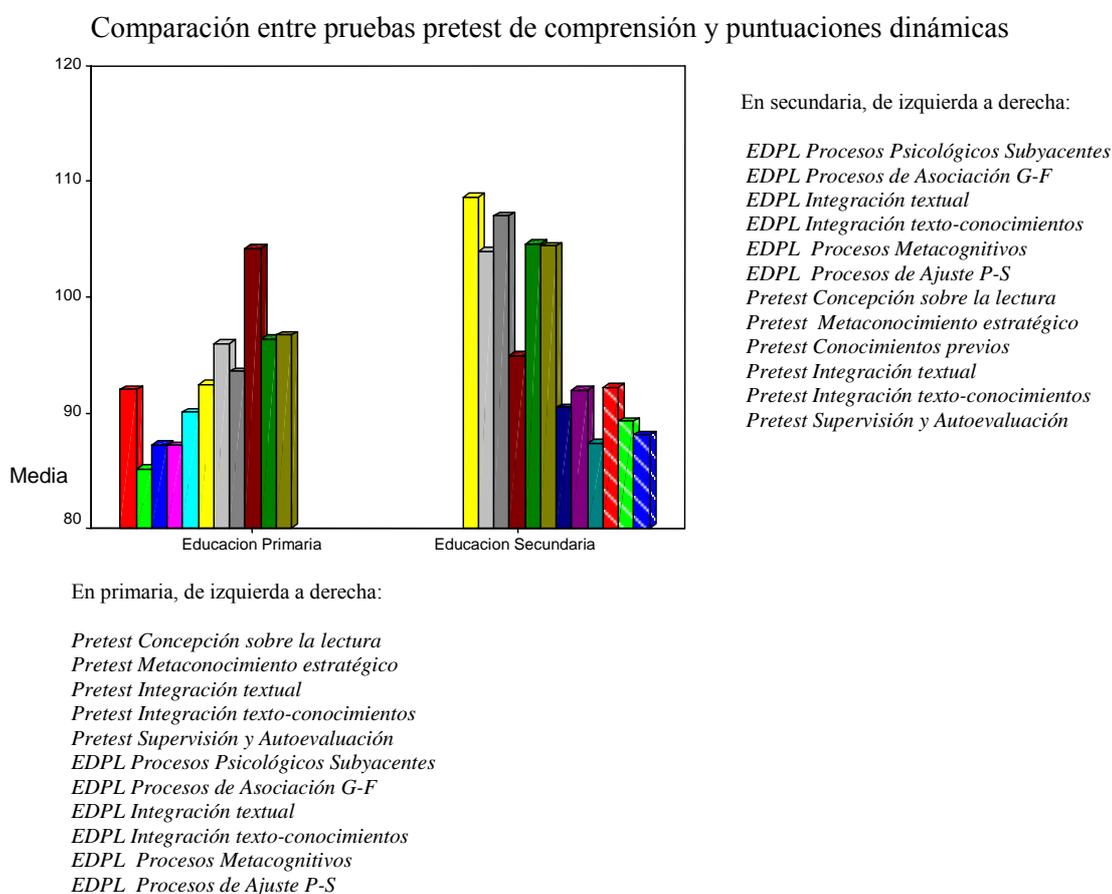
En primaria, la fase pretest de la prueba ECO 1 reveló que los procesos que habían representado una mayor dificultad fueron los de *Metaconocimiento estratégico*, seguido de los de *Integración texto-conocimientos previos* e *Integración textual*. En la fase postest, la dimensión *Integración textual* seguía siendo una de las que menor puntuación había obtenido, seguida de *Metaconocimiento estratégico* que, a pesar de haber experimentado una subida de dos puntos, seguía representando una mayor dificultad con relación al resto de dimensiones. En cuanto a las correlaciones, el análisis efectuado revela que el valor de correlación de la dimensión *Integración textual* con la puntuación dinámica relativa a estos procesos pasa de ser negativa en el pretest a tener un valor positivo en el postest ($r.0,189$). También la dimensión *Integración texto-conocimientos previos* pasa de un valor muy bajo de correlación en el pretest ($r.0,046$) a un nivel medio en el postest ($r.0,197$). Por su parte, la puntuación en la dimensión *Metaconocimiento estratégico* no mantiene ninguna correlación positiva con la puntuación dinámica relativa a estos procesos; los valores son negativos en el pretest, aunque no significativos, y, aunque de menor valor, también negativos con el postest de la prueba ECO 1. Idénticos resultados obtuvimos en los análisis correlacionales efectuados para la otra dimensión metacognitiva, *Supervisión y autoevaluación del proceso de comprensión*.

En síntesis, estos datos, en lo relativo a la localización de las mayores dificultades, apuntan en una dirección similar a la que podía derivarse del análisis del dispositivo de evaluación dinámica; sin embargo, más allá de la confirmación general de los resultados obtenidos, confirmación que se da sobre todo a nivel de *dónde* se encuentran algunas de las principales dificultades relacionadas con la lectura y la comprensión en sujetos con dificultades de aprendizaje, podría existir una escasa relación entre la información que puedan llegar a aportar pruebas estáticas y dinámicas con relación a determinados procesos implicados en la lectura; estas limitaciones observadas, sobre las que volveremos más tarde, estarían relacionadas con: (a) la capacidad diferencial de ambos modelos para profundizar y afinar en el análisis de *dónde* se encuentran las dificultades de lectura y de por qué se producen; (b) la capacidad diferencial para explicar el proceso mismo aprendizaje (integración de modelos de acción, incorporación de estrategias, etc.); (c) la capacidad diferencial para facilitar y explicar un mayor rendimiento en las actividades desarrolladas mediante la aplicación dinámica.

En secundaria, el análisis de las distintas dimensiones contempladas en la prueba ECOS, nos reveló que las de *Conocimientos previos*; *Supervisión del proceso de comprensión*; e *Integración texto-conocimientos previos*; fueron las que, a tenor de las puntuaciones promedio obtenidas por los sujetos experimentales, representaron una mayor dificultad en la fase pretest. En el postest todas estas puntuaciones se incrementaron significativamente, aunque la dimensión *Conocimientos previos* seguía figurando como la más difícil, esta vez junto a *Integración textual* e *Integración texto-conocimientos previos*. De nuevo aparecen como procesos especialmente problemáticos para el alumnado, aquellos relativos a la integración del texto con los conocimientos previos, así como a las habilidades metacognitivas, esta vez sólo en el pretest. Podemos comprobar que la evolución positiva de las puntuaciones en la prueba de comprensión para el grupo experimental de secundaria, hace desaparecer a los procesos metacognitivos de entre los más difíciles; el trabajo específico realizado desde las sesiones del dispositivo sobre estos procesos habría tenido un peso notable en este hecho, con una ganancia promedio para el GE tomado en su totalidad en torno a 10 puntos. Por otra parte, nuestra comprobación del comportamiento de las puntuaciones

de estas dimensiones en el GC evidencia que las ganancias postest son sensiblemente menores que las del GE; especialmente para la dimensión *Supervisión del proceso de comprensión*, la evolución de las puntuaciones fue significativamente menor (en torno a 2 puntos), no produciéndose diferencias significativas entre pretest y postest, ni el tránsito anterior desde la dificultad a la no dificultad. Si bien es cierto que no podemos olvidar que la puntuación inicial del GC era sensiblemente superior a la obtenida por el GE, con lo que quizás el margen de ganancia se pudo ver reducido, la puntuación final obtenida por este último grupo superó a la obtenida por el GC, aún partiendo de una posición inferior.

Figura 6. Comparación entre las puntuaciones obtenidas en las pruebas pretest de comprensión (ECO 1 y ECOS) y las puntuaciones dinámicas de los distintos procesos contemplados en el dispositivo



La inspección gráfica de los datos nos ofrece algunos elementos de interés. En primer lugar, tanto en primaria como en secundaria, la totalidad de las puntuaciones dinámicas se sitúa por encima de las puntuaciones obtenidas en los respectivos pretest de las pruebas de comprensión. A pesar de que ambas puntuaciones no son comparables de forma directa, el contraste efectuado nos ofrece una importante información. Ésta es relativa a que la valoración cuantificada del proceso de aprendizaje es más alta que la valoración obtenida de la pasación de las pruebas de evaluación de productos. Lógicamente, la aplicación dinámica del dispositivo EDPL contempla la mediación del aplicador en la consecución de las actividades propuestas con el fin de obtener el máximo de información sobre el desempeño y las dificultades del sujeto, y esta

mediación, como es sabido, conlleva en ocasiones la realización de la tarea en situaciones de colaboración. Sin embargo, sería un error bajo nuestro punto de vista pensar que este hecho resta valor al dato aportado anteriormente; como hemos tenido oportunidad de ver a lo largo de este trabajo, las situaciones de aprendizaje mediado nos dan una posibilidad idónea de observar y valorar en la *Zona de Desarrollo Próximo* que es capaz de realizar el sujeto con la ayuda ajustada. De este modo, la información obtenida y mostrada en el gráfico mostraría claramente que los sujetos habrían ejecutado significativamente mejor tareas de lectura y comprensión en situaciones de colaboración que en una situación de prueba estática e individual. Este resultado nos llevaría directamente a la propensión de que los mismos sujetos podrían llegar a realizar próximamente las mismas o similares actividades de forma autónoma. Como hemos tenido oportunidad de comprobar, esto se ha cumplido ya en parte, y de forma diferenciada en función de los distintos subgrupos, en el corto espacio de tiempo transcurrido en el estudio. Por otra parte, podemos comprobar en el gráfico que a pesar de los bajos promedios dinámicos obtenidos por el bloque de procesos de Integración texto-conocimientos previos, las puntuaciones transformadas, relativas al grupo de referencia, muestran diferencias notables con respecto a las obtenidas por el mismo grupo de sujetos en la dimensión de Integración texto-conocimientos en las pruebas estáticas. Esta diferencia es especialmente llamativa en primaria; sin embargo, en secundaria el promedio de estos procesos era sensiblemente menor y la puntuación transformada muestra una menor diferencia, aunque muy importante, entre ambas puntuaciones. También es destacable la diferencia en ambas etapas entre las dimensiones metacognitivas de la prueba criterio y la puntuación dinámica de los procesos metacognitivos.

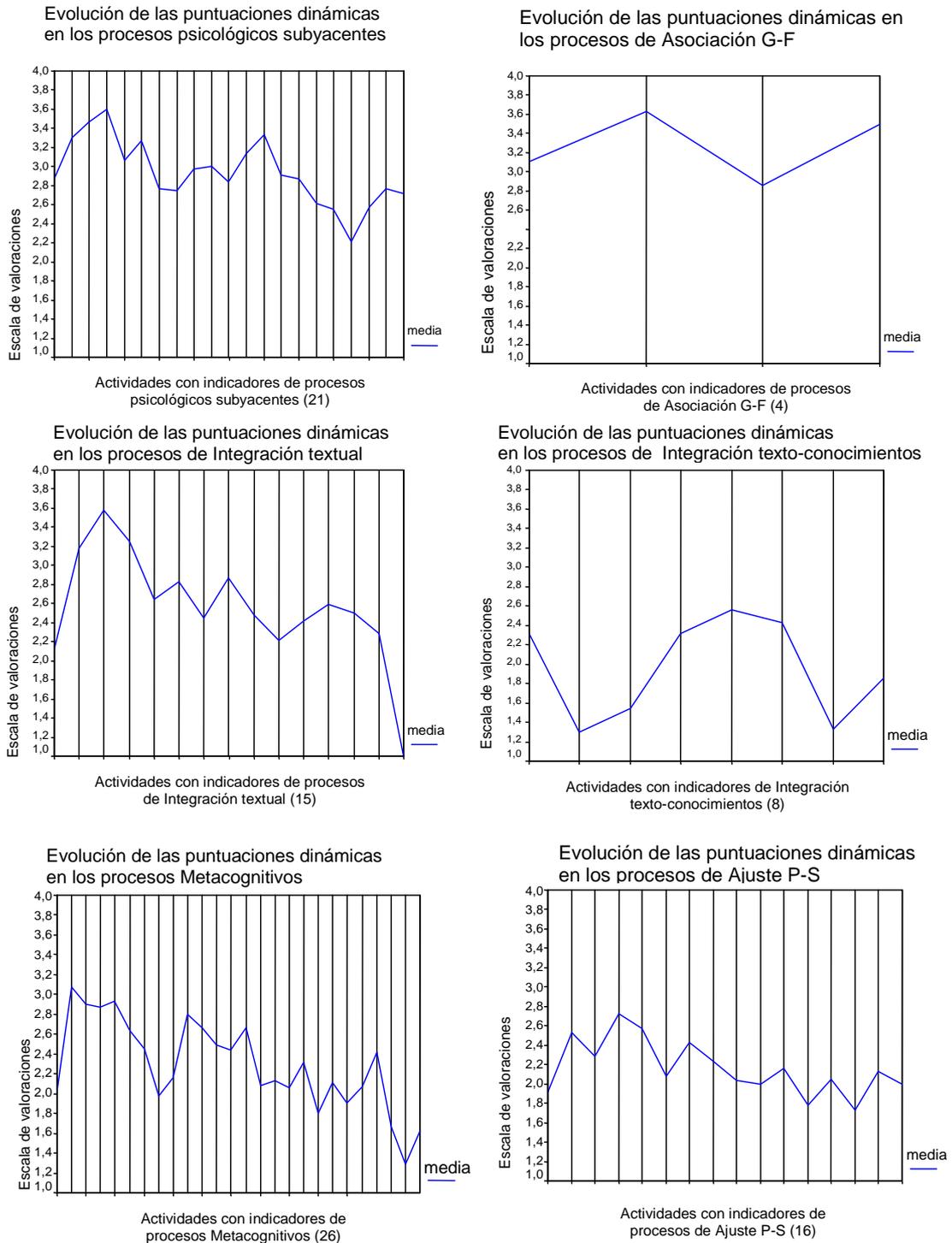
En cuanto a las correlaciones obtenidas, podemos destacar que la puntuación dinámica de los procesos de asociación G-F correlacionó positivamente y de forma muy significativa en el pretest con la dimensión *Integración texto-conocimientos previos* de la prueba criterio; así mismo, mantuvo una correlación significativa al nivel 0,05 con la dimensión *Conocimientos previos* y con la puntuación total de la prueba. En el postest los valores de correlación se incrementaron notablemente, estableciéndose una correlación muy significativa con las dimensiones *Conocimientos previos*, *Integración texto-conocimientos previos*, e igualmente con la puntuación total. Por otra parte, también mantuvo correlaciones significativas, aunque de menor nivel, con las dimensiones *Metaconocimiento estratégico*, *Integración textual*, y *Supervisión del proceso de comprensión*. Por su parte, la puntuación dinámica de los procesos de Integración textual mantenía una correlación significativa con la dimensión *Conocimientos previos*, aunque no con la dimensión relativa a los *procesos de Integración textual* en la prueba criterio; así mismo, se observó un incremento notable en la correlación obtenida en el postest con la misma dimensión *Conocimientos previos*. La puntuación dinámica de los procesos de integración texto-conocimientos previos mantuvo unas correlaciones muy bajas o incluso negativas con la mayoría de las dimensiones de ECOS, exceptuando la dimensión *Concepción sobre la lectura*, con la que correlacionó positivamente e incrementó su valor en la fase postest, llegando casi a la significatividad estadística ($r=0,378$; $p<0,062$). Por su parte, la puntuación dinámica de los procesos metacognitivos también correlacionó negativamente con la mayoría de las dimensiones de la prueba criterio de comprensión; estos valores se dieron tanto en el pretest como en el postest. Sólo para la dimensión *Concepción sobre la lectura*, al igual que con los procesos de integración texto-conocimientos previos, se obtuvo un valor medio y positivo de correlación ($r=0,252$), observándose igualmente un incremento

notable con respecto al valor de correlación obtenido en el pretest. Por último, la puntuación dinámica de los procesos de ajuste personal-social también obtuvo valores negativos de correlación con la mayoría de las dimensiones de ECOS, en pretest y en postest. De nuevo, sólo la dimensión *Concepción sobre la lectura* mantuvo una correlación positiva, incrementando notablemente su valor en el postest, aunque sin llegar a la significatividad estadística ($r=0,265$; $p<0,200$). En definitiva, hemos podido comprobar que los valores de correlación encontrados entre las puntuaciones dinámicas de los distintos procesos contemplados en el dispositivo EDPL y las distintas dimensiones consideradas en las pruebas criterio de comprensión, son en general de un nivel medio o bajo, cuando no negativos. Salvo la puntuación dinámica de los procesos de asociación G-F, que mantiene correlaciones positivas y significativas con varias de las dimensiones, el resto de procesos y especialmente los procesos metacognitivos, los de integración texto-conocimientos previos, así como los de ajuste personal-social, muestran valores de correlación muy bajos o negativos. Los dos primeros resultaron ser los más difíciles, tanto en la aplicación dinámica, como, en buena medida, también en las pruebas de comprensión, tanto en primaria como en secundaria, lo que, sin embargo, no parece haber contribuido a una mayor relación entre las puntuaciones. Tal y como apuntábamos anteriormente, estos resultados parecen demostrar que las puntuaciones extraídas de la valoración de un proceso de aprendizaje, en este caso el proceso de lectura y comprensión, poco tendrían que ver con las puntuaciones obtenidas de la valoración del producto de aprendizaje.

Evolución de las puntuaciones dinámicas en el transcurso de la aplicación de EDPL

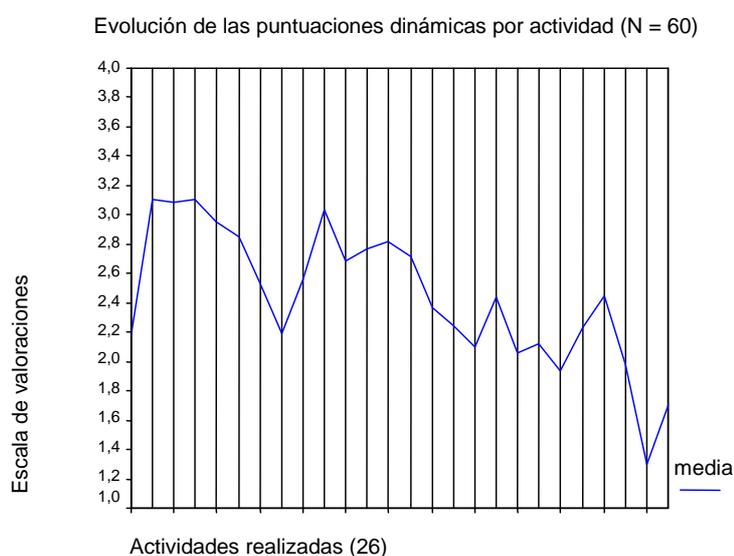
Un aspecto de especial interés para nosotros es el que tiene que ver con el desarrollo evolutivo de las puntuaciones dinámicas durante la aplicación del dispositivo. El análisis de esta evolución podría ser reflejo del proceso de aprendizaje de los sujetos acerca de aquellas habilidades trabajadas, así como del proceso de mediación efectuado por los aplicadores. A continuación mostramos gráficamente cómo se ha producido la evolución de las puntuaciones dinámicas para cada uno de los procesos contemplados en el dispositivo EDPL a lo largo de la fase experimental de nuestro estudio, junto al número de actividades que contaban con indicadores de cada uno de estos procesos. En estos gráficos podemos observar cómo se ha producido la evolución de las puntuaciones dinámicas para el grupo experimental ($N = 60$) en cada uno de los procesos contemplados en el dispositivo EDPL, a lo largo del desarrollo de la aplicación. La inspección visual de los datos de esta forma dispuestos, nos ofrece la posibilidad de observar la valoración del proceso de trabajo, y así mismo la tendencia en las puntuaciones otorgadas con base en las valoraciones del profesorado aplicador. De este modo, en varios procesos podemos observar una cierta tendencia descendente en los promedios obtenidos a medida que avanzaba la aplicación. Este descenso es especialmente notable en el bloque de procesos de integración textual y en los procesos metacognitivos, y más atenuada en los procesos psicológicos subyacentes y en ajuste personal-social. Por su parte, la tendencia observada para el bloque de procesos de asociación G-F, a través de un número significativamente menor de actividades en las que éstos podían valorarse, expresa cambios ascendentes y ascendentes no significativos, manteniéndose en cualquier caso en un nivel alto de la escala de valoración. Los procesos de integración texto-conocimientos previos, cuya línea de evolución se sitúa significativamente por debajo de las demás, observan una tendencia parecida a la anteriormente mencionada en cuanto a las subidas y bajadas, aunque con picos mucho más acusados.

Figura 7. Evolución de las puntuaciones dinámicas en el transcurso de la aplicación de EDPL



También podemos observar una tendencia descendente en la gráfica que muestra la evolución del promedio de las puntuaciones dinámicas globales, obtenidas en cada una de las actividades realizadas. Es preciso volver a resaltar aquí las diferencias existentes en cuanto al número y al tipo de actividades puestas en práctica en función de los diversos grupos de aplicación, lo que sin duda ha influido notablemente en la configuración de la gráfica. En general, podemos comprobar la tendencia apuntada anteriormente, relativa a una mayor dificultad encontrada por los sujetos a medida que se avanzaba en la aplicación del dispositivo. Estas mayores dificultades se habrían traducido en valoraciones más pobres sobre la puesta en práctica de los indicadores propuestos y, por ende, en puntuaciones menores.

Figura 8. Evolución de las puntuaciones dinámicas obtenidas en las actividades del dispositivo EDPL



Relaciones entre la necesidad de mediación y la puntuación final obtenida

Apuntábamos en páginas anteriores la doble dirección contrapuesta que podía tener la hipótesis que relacionaba la variable considerada en la matriz de datos relativa a la necesidad de mediación de los sujetos con la puntuación dinámica obtenida en EDPL. Esta doble dirección se origina en que la interpretación de los comentarios realizados por los aplicadores podía efectuarse en un doble sentido. Por una parte, las valoraciones hacían referencia a que un determinado sujeto *se había beneficiado* de la mediación en la puesta en práctica de un determinado indicador en una tarea; por otra, las valoraciones aludían a la *necesidad de mediación* para que el sujeto pudiera acceder a la puesta en práctica de ese indicador y, por tanto, a la ejecución de la tarea. Sin ser contradictorias, esta doble interpretación podía dar pie a esperar distintos resultados en los análisis de correlación efectuados. La primera idea podría hacer pensar en correlaciones positivas entre el número de veces que el alumno se benefició de la mediación y la puntuación final obtenida; la segunda, por el contrario, podría llevarnos a una correlación negativa entre el número de veces que los sujetos necesitaron la mediación del aplicador para solucionar o avanzar en las tareas y la puntuación final. Los resultados los mostramos en la tabla 74.

Tabla 74. Correlaciones entre el número de veces que los sujetos necesitaron/se beneficiaron de la mediación del aplicador y las puntuaciones dinámicas obtenidas en el dispositivo

	Procesos Psicológicos Subyacentes	Asociación Grafema-Fonema	Integración textual	Integración texto-conocimientos	Procesos Meta-cognitivos	Ajuste personal-social	Puntuación dinámica global
Mediación (media=20,52)	-0,527**	-0,548**	-0,460**	0,242	-0,332**	-0,135	-0,410**

Como podemos comprobar, los resultados del análisis efectuado no parecen dejar lugar a muchas dudas acerca de qué interpretación de las que mencionábamos anteriormente era la más ajustada. Independientemente de que los sujetos se beneficiaran de la mediación ejercida por los aplicadores en el transcurso de las tareas, como hicieron notar los aplicadores en las sesiones del Grupo de Trabajo, parece claro que las valoraciones del grupo de aplicadores sobre la respuesta a la mediación, expresadas concretamente en referencia a indicadores determinados en el transcurso de una actividad, hacían alusión sobre todo a la necesidad de mediación de los sujetos para acceder a, o mejorar, la puesta en práctica de determinadas habilidades evaluadas mediante los indicadores propuestos. Salvo los procesos de integración texto-conocimientos, todos los procesos, así como la puntuación dinámica global, correlacionan negativamente con el número de veces que el sujeto necesitó la mediación del aplicador; es decir, las puntuaciones dinámicas obtenidas fueron más altas en aquellos sujetos que requirieron menor cantidad de intervenciones mediadoras. El resultado observado en los procesos de integración texto-conocimientos podría relacionarse con la mayor dificultad que estos procesos supusieron para el alumnado. En páginas anteriores comprobamos que, junto con los procesos metacognitivos y los procesos de ajuste personal-social, los de integración texto-conocimientos fueron los que representaron una mayor dificultad, obteniendo significativamente una menor puntuación dinámica con relación al resto de procesos. Precisamente estos tres bloques de procesos son los que obtienen valores de correlación menos negativos. Los tres procesos restantes, sin embargo, que obtuvieron mayores puntuaciones dinámicas, correlacionan negativa y muy significativamente, con valores más altos. En efecto, en la medida en que los indicadores de evaluación pertenecen a procesos que resultaron difíciles para los sujetos, la necesidad de mediación habría sido más lógica y esperable, y por tanto menos indicativa de dificultades en los sujetos. Por el contrario, la necesidad de mediación en los indicadores de procesos que representaron un menor problema para la mayoría de los sujetos, habría mostrado la presencia de mayores dificultades en los sujetos.

Los resultados del dispositivo EDPL en función de los grupos de aplicación

Hemos querido por último mostrar los resultados obtenidos en el dispositivo EDPL en función de los distintos grupos de aplicación. El análisis de estos resultados nos será especialmente útil posteriormente, cuando abordemos las relaciones entre los patrones metodológicos de evaluación dinámica puestos en práctica por los aplicadores y los resultados obtenidos en la aplicación del dispositivo. La tabla que exponemos a continuación muestra la puntuación obtenida en cada uno de los procesos contemplados, así como la puntuación dinámica global para cada uno de los grupos participantes.

Tabla 75. Puntuaciones dinámicas obtenidas en cada uno de los grupos de aplicación

Grupos de aplicación	Procesos Psicológicos Subyacentes	Asociación Grafema-Fonema	Integración textual	Integración texto-conocimientos	Procesos Meta-cognitivos	Ajuste personal-social	Puntuación dinámica global
(A) CEIP Andalucía	107,196	111,333	102,820	102,937	108,953	103,315	108,936
(B) CEIP A. Machado	81,008	90,017	89,903	104,617	88,439	85,354	85,326
(C) CEIP El Ruedo	71,273	85,878	79,334	91,554	79,792	78,654	75,519
(D) CEIP El Ruedo	90,517	79,670	82,249	90,090	74,648	72,309	74,082
(E) CEIP Lope de Rueda	72,111	101,399	86,612	97,557	74,017	79,848	73,779
(F) CEIP Sánchez Alonso	98,538	87,172	95,437	116,395	106,523	116,010	105,668
(G) IES Leonor de Guzmán	103,438	112,782	105,993	88,895	92,942	94,098	96,583
(H) IES Leonor de Guzmán	112,648	95,967	106,523	93,782	101,452	104,149	101,285
(I) IES Mariana Pineda	105,506	102,020	103,624	86,182	94,011	95,356	95,790
(J) IES Torreblanca	114,008	92,500	106,952	124,479	122,990	122,479	121,887
(K) IES Torreblanca	112,451	107,918	112,675	85,703	120,511	114,242	120,341

Análisis del dispositivo en función del perfil de cuatro puntos

INFORMACIÓN COMPARATIVA DEL DISPOSITIVO EDPL

En primer lugar nos dirigimos a valorar diferencialmente el modelo propuesto con relación a otros modelos dinámicos de los que hemos descrito y analizado en nuestro trabajo. En este sentido, ya hemos hecho alusión a que la aplicación del dispositivo EDPL no suponía en ningún caso una instrucción centrada en una determinada prueba que posteriormente sería evaluada. Tampoco en el contexto de nuestro estudio esto ha tenido lugar, ya que las tareas desarrolladas mediante la aplicación dinámica no significaban un entrenamiento para ninguna de las pruebas de criterio utilizadas. Este hecho supone en la práctica una distinción con algunos procedimientos dinámicos de evaluación en los que, como hemos tenido oportunidad de comprobar en este trabajo, el contenido de la fase de instrucción es sólo ligeramente diferente al contenido de las pruebas criterio. En estos casos, el objetivo de la evaluación dinámica es sobre todo demostrar que la mediación de un determinado contenido de aprendizaje mejora la actuación del sujeto en tareas de rendimiento similares a las entrenadas. Aceptando la pertinencia y el interés que puede tener comprobar empíricamente este hecho, lo cierto es que presenta algunos problemas, sobre todo desde una perspectiva aplicada. En efecto, la excesiva similitud de algunas de las tareas entrenadas con las que posteriormente serán evaluadas, resta potencia al modelo en cuanto a su capacidad de transferencia; es decir, el modelo nos plantea dudas acerca de que las mejoras observadas en los sujetos puedan ser atribuidas de forma significativa a ganancias efectivas en el constructo evaluado, ya que podrían más bien ser expresión del aprendizaje del sujeto en unas tareas determinadas y siguiendo un determinado procedimiento válido para la resolución de esas tareas.

Por otra parte, también hemos analizado críticamente el uso masivo que se ha realizado en la evaluación dinámica de las pruebas denominadas «libres de cultura», tales como el test de Raven o el test de factor «G» de Cattell. Expusimos que las razones principales para ello las encontrábamos en que al no presentar contenidos de tipo cultural y que, por lo tanto, pudieran establecer diferencias de partida entre los sujetos evaluados, se prestaban mejor a la filosofía de la evaluación dinámica, a través de la cual se quería valorar el potencial de aprendizaje en condiciones de igualdad para todos los sujetos. Diversas investigaciones en cambio han mostrado de una u otra forma que las condiciones de partida no se igualan por el hecho de utilizar pruebas sin contenidos de tipo cultural (Cole y otros, 1971; Cole y Scribner, 1974; Laboratory of Comparative Human Cognition, 1982; Sternberg y otros, 1996; Sternberg y otros, 1999; Sternberg y Grigorenko, 2003), ya que aspectos que acompañan habitualmente a la pasación de las pruebas como: (a) la experiencia previa en situaciones de prueba, (b) el mayor contacto y conocimiento de los sujetos favorecidos socialmente sobre las actividades con contenidos similares a las propuestas en los test, así como (c) el mayor significado social que tiene para algunos sujetos el contenido y la propia actuación en pruebas de este tipo; pueden hacer que el rendimiento efectivo de los sujetos pueda verse afectado notablemente. Si esto es así, las pruebas dinámicas podrían utilizar en mayor medida actividades que estuvieran más relacionadas con contenidos de tipo cultural que estuvieran presentes en el contexto social y educativo. Las ventajas que esto aportaría a la investigación lo son desde una doble perspectiva; por una parte, los resultados ganarían en validez ecológica, ya que los contenidos trabajados durante la fase de intervención y evaluación podrían relacionarse con los contenidos escolares. Esto disminuiría los problemas de transferencia antes mencionados y facilitaría efectuar predicciones sobre el rendimiento escolar que estuviesen mejor fundadas. Por otra parte, también podría valorarse el *proceso de aprendizaje* del sujeto en actividades en las que encuentra dificultades, deduciendo de la aplicación dinámica aquellas pautas metodológicas que han demostrado ser más eficaces en la construcción de conocimientos por parte del sujeto.

Otro aspecto relevante con relación a la evaluación diferencial, sería el que hace referencia a las divergencias que podríamos establecer con pruebas convencionales que evalúan el mismo constructo. Estas podrían valorarse con relación a tres indicadores: (a) la calidad de los datos obtenidos, (b) el tipo de información que ofrece y (c) las aportaciones metodológicas suplementarias a la información obtenida con la pasación de pruebas estáticas.

Calidad de los datos obtenidos. La fiabilidad de las valoraciones en EDPL

En primer lugar, con relación a la calidad de los datos, debemos hacer hincapié junto con Sternberg y Grigorenko (2003), en que, más allá de los parámetros estadísticos de coherencia interna del contenido de las tareas propuestas, aquello que constituye verdaderamente un reto para las pruebas dinámicas es la fiabilidad que puedan mostrar ante la valoración que efectúen distintos evaluadores sobre los aspectos centrales que se contemplan. Esta valoración, en el caso del dispositivo EDPL, debía realizarse fundamentalmente sobre: (a) la puesta en práctica de los procesos implicados en cada una de las tareas, especialmente de los procesos metacognitivos; (b) el proceso seguido por el sujeto para realizar o resolver la actividad propuesta; y (c) el propio proceso de mediación efectuada. A un tiempo, esta valoración informaría, a través de los indicadores o criterios de evaluación propuestos para cada una de las actividades, acerca

de dónde se encontraban principalmente las dificultades del sujeto. Por último, el objetivo del proceso de evaluación dinámica propuesto era utilizar la información precedente para la optimización del proceso de aprendizaje y, especialmente, del proceso de comprensión. El análisis de las valoraciones cualitativas efectuadas por los aplicadores sobre algunos de los aspectos anotados anteriormente, será efectuado en un apartado posterior. En este, nos centraremos en la valoración cuantitativa efectuada sobre los indicadores de evaluación de cada uno de los procesos evaluados, así como en algunos aspectos cualitativos relativos a ésta. En este sentido, tal y como relatábamos cuando hacíamos alusión al procedimiento seguido en la investigación, la utilización del video en las sesiones del Grupo de Trabajo, fue un elemento que se tornó clave en el proceso de formación continua de los aplicadores, así como en la continua y difícil búsqueda de la unificación de criterios para valorar los indicadores de evaluación de las actividades. Uno de los objetivos principales, por tanto, del uso del video en estas sesiones era el de lograr el máximo acuerdo entre los distintos observadores y aplicadores del dispositivo EDPL a la hora de realizar una valoración sobre la puesta en práctica de un determinado indicador. Lamentablemente, como ya apuntamos en páginas anteriores, no se ha podido realizar un estudio, con psicólogos ajenos al proceso de investigación, para evaluar la fiabilidad en el acuerdo inter-observadores.

A pesar de las dificultades prácticas que la aplicación de un dispositivo como el que presentamos, ofrece para calcular índices de fiabilidad en función de la valoración diferencial de los distintos aplicadores, hemos tratado de realizar una estimación de este parámetro para EDPL con relación a las distintas valoraciones efectuadas. Para ello hemos debido seleccionar necesariamente aquellas actividades que fueron realizadas por un mayor número de sujetos y hemos calculado posteriormente el coeficiente Alfa para las valoraciones dinámicas realizadas por los aplicadores sobre los indicadores de evaluación propuestos en EDPL para cada uno de los procesos contemplados. Las actividades seleccionadas (en realidad, la consideración de otras actividades no permitía la realización de los cálculos, debido al reducido número de casos para efectuar las relaciones) contenían indicadores pertenecientes a todos los procesos contemplados en el dispositivo: 13 indicadores relativos a los procesos psicológicos subyacentes, 4 relativos a los procesos de asociación G-F, 25 a los procesos de integración textual, 18 a los procesos de integración texto-conocimientos previos, 71 a los procesos metacognitivos y 18 a los procesos de ajuste P-S. Las actividades consideradas fueron finalmente nueve: (10) Lectura de pseudopalabras; (13) Lectura de palabras; (16) Ordenar frases; (17) Relacionar frases-dibujos; (18) Elaborar frases nuevas; (19) Signos de puntuación; (22) Uso de la estructura de un texto; (23) Extracción y construcción del significado; (24) Integración del texto e inferencias.

El análisis se ha efectuado sobre un total de 15 casos, siendo el número de ítems (indicadores valorados por los aplicadores en las 9 actividades) de 149. Por tanto, el número total de observaciones sobre las que se han efectuado estos cálculos es de 2.235. Los resultados globales muestran un valor Alfa de 0,98, siendo la correlación inter-ítems de 0,26. Por su parte, el contraste efectuado mediante el análisis de varianza resultó muy significativo, con un valor F de 7,459 y una significación asociada de $p < 0,000$. Por otra parte, tal y como apuntábamos, hemos calculado también el coeficiente de correlación intra-clase. Este coeficiente permite conocer el grado de consistencia interna de los datos obtenidos, ofreciendo una estimación de la fiabilidad inter-observadores. Tanto el contraste para la medición simple (0,29), como el efectuado para la medición promedio del coeficiente (0,98), resultan estadísticamente muy

significativos ($F = 60,813$; $p < 0,000$ para 14 grados de libertad). La definición de la correlación intra-clase efectuada mediante la estimación del acuerdo absoluto entre las observaciones, muestra pequeñas variaciones sobre los datos anteriores. En este caso, la medición simple de la correlación intra-clase, se situó en 0,22, mientras que la medición promedio lo hizo igualmente en 0,98. El coeficiente de fiabilidad global de la escala fue también de 0,98, y los valores del contraste efectuado mostraron una F y una probabilidad asociada idénticas a los obtenidos en la opción anterior.

Hemos efectuado estos mismos cálculos para los ítems pertenecientes al bloque de procesos metacognitivos. Los resultados para estos 71 ítems muestran una correlación inter-ítems de 0,34 y un valor Alfa global de 0,97. El contraste para el análisis de varianza es muy significativo ($F = 5,309$; $p < 0,000$). Por su parte, el análisis de la consistencia interna mediante la estimación del coeficiente de correlación intra-clase, muestra en su medición simple un valor de 0,34 y en su medición promedio un valor de 0,97. El valor F del contraste efectuado es de 37,648 y la probabilidad asociada de $p < 0,000$. La estimación del acuerdo inter-observadores indica un valor en la medición simple de 0,29 y de 0,97 para la medición promedio. Los valores para la F y la probabilidad son idénticos que en la opción anterior, al igual que el valor Alfa global.

Por tanto, en función de los resultados obtenidos, podemos estimar que el nivel de consistencia interna de las observaciones efectuadas, ha sido muy alto, obteniéndose valores estimados de la fiabilidad inter-observadores que se acercan al acuerdo absoluto, tanto en la estimación efectuada con la totalidad de los indicadores valorados en la actividades, como en la realizada con los indicadores de procesos metacognitivos.

Otro aspecto relativo a la calidad de los datos que puedan obtenerse, es el que guarda relación con la tenencia en cuenta de los grupos control. Se ha hecho alusión en numerosas ocasiones, y por parte de distintos investigadores (Mora, 1998; Sternberg y Grigorenko, 2003), a la problemática que acompaña habitualmente a la, por otra parte, necesaria consideración de grupos de control en los estudios. Se ha destacado que muchas investigaciones relacionadas con pruebas dinámicas adolecen de controles adecuados, ya que, contemplando asiduamente un grupo control de *no intervención*, carecen del control necesario sobre la variable *otros tratamientos*. Los argumentos empleados para cuestionar las conclusiones obtenidas sin la existencia de estos grupos *placebo*, hacen referencia a que no se puede tener certeza de que los resultados no se deban al factor *enseñanza*, independientemente del contenido de la propuesta trabajada. Por tanto, paradójicamente, un grupo control de no tratamiento añadiría problemas metodológicos a la investigación dinámica. En nuestro caso, ya apuntamos que los sujetos que no recibían EDPL, es decir, los sujetos del grupo control, recibían otros tratamientos. Esto es así debido a la propia lógica de la investigación en el contexto escolar. Los alumnos del grupo control no dejan de recibir clases en sus grupos de referencia, ni de recibir los distintos tipos de apoyo considerados. En los grupos de apoyo, estos sujetos participan de trabajos más o menos específicos y adecuados a sus necesidades educativas, y se benefician de intervenciones educativas que conllevan habitualmente periodos intensos de interacción con sus profesores. Por tanto, en nuestro estudio, el grupo de *no tratamiento* recibiría un nombre más adecuado si pasara a denominarse grupo de *otros tratamientos*. Por otra parte, ya comentamos que al tratarse de distintos tratamientos, realizados por distintos profesores y en distintos centros, los posibles efectos se suponen razonablemente aleatorizados, por lo que no cabría esperar influencia significativa en los resultados.

El tipo de información ofrecida y las aportaciones metodológicas adicionales

En cuanto al tipo de información que ofrece la aplicación del dispositivo de evaluación dinámica EDPL, así como a las aportaciones metodológicas diferenciales con relación a las pruebas estáticas, cabría preguntarse lo siguiente: ¿en qué medida los resultados obtenidos con la aplicación del dispositivo EDPL conducen a propuestas de intervención más eficaces para los sujetos evaluados? Sin duda esta cuestión, relacionada con la denominada *validez del tratamiento*, es clave para nosotros. En efecto, tanto en la teoría de base formulada, como en diferentes momentos a lo largo del presente trabajo, hemos hecho referencia a que el objetivo fundamental de la aplicación del dispositivo en el contexto escolar es la obtención de información útil para la mejora de los procesos de aprendizaje del sujeto; lógicamente, para que esta mejora de los procesos de aprendizaje tenga lugar, el proceso de evaluación dinámica debe conducir hacia una optimización de la acción educativa encaminada a la mejora del sujeto. Por tanto, el proceso de aplicación conlleva que el evaluador extraiga aquellos patrones metodológicos que se han revelado más eficaces, bien para la obtención de información sobre los procesos afectados en el sujeto, bien para la obtención de mejores resultados por parte del sujeto en el desarrollo de las actividades. Este proceso de *extracción o deducción* de los patrones metodológicos más eficaces queda facilitado en nuestra propuesta por la presencia en cada una de las tareas de los indicadores de evaluación. De esta forma, en función de la valoración que efectúa el evaluador de la puesta en práctica de los indicadores contemplados en cada una de las tareas realizadas, podría predecirse la respuesta de los sujetos a una posterior intervención educativa. Sin embargo, debido a que, como hemos ya referido, ninguno de los grupos de aplicación logró realizar todas las actividades propuestas para la investigación y, como consecuencia, no se ha experimentado una acción educativa posterior basada en los procesos trabajados y en los resultados obtenidos que fueran indicativos de mejora, no estamos aún en condiciones de confirmar empíricamente que la información extraída de EDPL conduce a una intervención más eficaz. Sí podremos, en cambio, hacer referencia posteriormente a la opinión de los profesores aplicadores sobre esta posibilidad, basándose en la experiencia de la fase experimental.

Por otra parte, como queda reflejado en el capítulo dedicado al análisis del dispositivo, la evaluación dinámica propuesta es en parte intervención, por lo que tendríamos la oportunidad de valorar, aún parcialmente, en qué medida mejora la actuación del sujeto en aquellos procesos que el aplicador valoraba desde el inicio como claves en la explicación de las dificultades, o bien a través del uso de patrones metodológicos que se revelaban claves en la explicación de su mejora durante el desarrollo de las actividades. En línea con la argumentación anterior, también podrían utilizarse las valoraciones realizadas por los aplicadores, es decir, las puntuaciones obtenidas de la aplicación del dispositivo, para predecir la actuación diferencial de los sujetos en el postest, entendiendo que éste reflejaría la «enseñanza» obtenida.

Cuando abordábamos algunos de los problemas planteados en la evaluación de la comprensión de textos, dentro del capítulo dedicado al análisis del dispositivo EDPL, exponíamos las razones que a nuestro juicio avalaban un necesario cambio en la perspectiva de evaluación. Al mismo tiempo planteábamos algunas propuestas que a nuestro entender podrían optimizar el proceso de evaluación. Los problemas podrían sintetizarse básicamente en: (a) la ejecución individual de las pruebas por parte de los sujetos, (b) la valoración exclusiva del producto de la comprensión, (c) la

descontextualización de los instrumentos, y (d) la ausencia de evaluación del contexto. Por otra parte, nuestras propuestas hacían referencia a la evaluación interactiva del proceso de resolución de la actividad de comprensión, tratando de acceder al conocimiento consciente de las propias estrategias y de valorar procesos metacognitivos de control y regulación de la actividad de comprensión durante la propia actividad. La evaluación dinámica del proceso de comprensión y de los procesos implicados, otorgaría información más valiosa sobre las dificultades del sujeto al afrontar tareas de lectura, al tiempo que facilitaría el conocimiento del evaluador sobre cómo ayudar mejor al sujeto mediante la extracción de patrones metodológicos que se hubieran mostrado eficaces durante el proceso de aplicación. Igualmente proponíamos la contextualización de los instrumentos de evaluación y la evaluación del contexto de aula dentro del proceso de evaluación.

Por tanto, la información obtenida a través de la aplicación del dispositivo EDPL es cualitativamente distinta a la que podemos obtener con la pasación de una prueba estática que valore la comprensión. Esta distinción la podemos efectuar también con relación a nuestras propias pruebas de comprensión, elaboradas para esta investigación. Aunque, como dijimos, estas pruebas fueron construidas tratando de responder de la forma más completa posible a la valoración de los procesos implicados en la comprensión, hemos sido conscientes en todo momento de que se trataba finalmente de pruebas estáticas, que informaban fundamentalmente sobre el producto de la comprensión.

Por último, para cerrar este epígrafe en torno a la validez diferencial del dispositivo, un índice que podría ser revelador de la *validez de criterio* en EDPL sería el que correlacionara los resultados obtenidos tras la aplicación del dispositivo, bien con las puntuaciones postest obtenidas por los sujetos, bien con un indicador del rendimiento académico. A pesar de que todos los sujetos del GE presentaban dificultades de aprendizaje de la lectura, podrían observarse agrupamientos en función de los resultados obtenidos que validaran el dispositivo en este sentido. Este planteamiento nos conduce directamente al siguiente punto considerado en nuestro esquema, es decir, el poder de predicción.

EL VALOR PREDICTIVO DE EDPL

Ya hemos cuestionado en este trabajo la excesiva atención que durante largo tiempo se prestó en la investigación dinámica al valor predictivo de las pruebas (Sternberg y Grigorenko, 2003), centrando casi de forma exclusiva la investigación en el hallazgo de valores que predijeran comparativamente mejor que las pruebas estáticas de CI, el rendimiento posterior de los sujetos. Dado que en los estudios que han incidido primordialmente en la predicción, no se ha considerado habitualmente el rendimiento académico como un factor idóneo para evaluar la validez de las pruebas, el problema ante esta perspectiva podría ser «quedarse» exclusivamente en el análisis de un determinado parámetro estadístico —en este caso la validez—, sin que ello signifique repercusión alguna en el contexto para el que se supone está ideada la prueba. En este trabajo, nos hemos hecho eco de propuestas alternativas de evaluación dinámica para las que la predicción no era el objetivo fundamental (Delclos, Vye, Burns, Bransford y Hasselbring, 1992; Ruijsenaars, Castelijns y Hamers, 1993; Jensen, 2000). Estas propuestas entienden que el objetivo básico es la mejora de los sujetos y, por tanto, que las pruebas dinámicas deben orientarse fundamentalmente hacia la intervención. En

línea con estos postulados, nosotros hemos hecho referencia también a que la relevancia de analizar y establecer el valor predictivo de una prueba dinámica, pasa porque este análisis sea efectuado con el fin de obtener datos que permitan realizar propuestas de acción educativa que optimicen el proceso de aprendizaje de los sujetos.

A nuestro entender, en la evaluación dinámica el análisis del valor de predicción debería mostrar dos tipos de índices; por una parte, debemos establecer el valor predictivo en función de los *resultados*; por otra, deberíamos poder establecer el valor de predicción del *proceso de aplicación* de la prueba dinámica. En cuanto al primero de ellos, éste vendría dado, como es habitual, por la correlación existente entre la puntuación dinámica —ya se obtenga ésta a través del cálculo de la diferencia entre las puntuaciones pretest y postest, o a través de la puntuación obtenida en las mismas pruebas dinámicas—, y un determinado valor considerado criterio para medir la validez del modelo. Este factor puede ser la puntuación obtenida en el mismo postest, en otras pruebas consideradas equivalentes a las utilizadas en el postest, o en valores procedentes del contexto de aplicación de la prueba, como pueden ser valores cuantificables del rendimiento académico. Por otra parte, el valor de predicción del proceso de aplicación dinámica, vendría dado por la relación entre aquellos patrones metodológicos que optimizaron el funcionamiento cognitivo o el ajuste personal-social del sujeto durante el desarrollo de la evaluación dinámica, y un valor procedente del contexto de aplicación de la prueba. Este valor de contexto —por ejemplo, el rendimiento académico— podría ser a su vez un valor cuantificable del resultado en pruebas criterio, o bien un valor, cuantitativo o cualitativo, obtenido del *proceso de aprendizaje* del sujeto. El establecimiento de este valor predictivo del proceso de evaluación dinámica sería, de este modo, similar al hecho de determinar los *mecanismos* a través de los cuales un proceso de intervención educativa ha tenido éxito, provocando mejoras significativas en los sujetos.

En la práctica educativa, la relación establecida para determinar el valor de predicción del proceso de evaluación dinámica, puede llevarse a cabo mediante la constatación de un evaluador que está inmerso en el contexto de intervención. El evaluador puede establecer mediante el análisis del caso, una relación entre aquellos patrones metodológicos que resultaron claves en la mejora del sujeto durante la aplicación dinámica, y su posterior actuación en procesos de intervención educativa. Por otra parte, en procesos de investigación, es posible establecer igualmente esta relación mediante la constatación «externa» de que la puesta en práctica de las propuestas o recomendaciones psicopedagógicas basadas en el proceso de aplicación, conduce a mejoras significativas del sujeto evaluado. De otra forma, puede también determinarse esta relación a nivel más cuantitativo. Es necesario para ello la asignación previa de valores cuantificables a indicadores de evaluación que expresen las habilidades requeridas, en función de la calidad de su puesta en práctica; posteriormente, debemos determinar aquellos patrones metodológicos utilizados durante el proceso de aplicación dinámica que se relacionan con mejores puntuaciones en los indicadores de evaluación. La determinación de estos patrones metodológicos facilitadores del acceso a los criterios de aprendizaje en el proceso de aplicación de la prueba dinámica, servirá para contrastar su eficacia en un proceso posterior de intervención educativa, en el que también deberemos lógicamente utilizar y dar valor a la puesta en práctica de los indicadores de evaluación mencionados como criterio del contexto de aplicación.

Valor de predicción con relación al rendimiento académico

Ya expusimos, cuando analizábamos las bases teóricas y metodológicas de nuestra propuesta, algunos de los problemas que los investigadores habían descrito con relación al uso del criterio *rendimiento académico* para evaluar la validez de los modelos dinámicos. Especialmente desde los modelos propuestos por Feuerstein y Budoff, se hacía hincapié en la falta de idoneidad de este criterio externo para validar sus pruebas. Las razones, esgrimidas y argumentadas principalmente por Lidz (1991), hacían referencia a que los modelos dinámicos se centran en los procesos de aprendizaje, mientras que los valores habituales de rendimiento académico son valores centrados en el producto del aprendizaje. Se ha apuntado igualmente desde esta perspectiva que la adecuación de un criterio externo de la validez de las pruebas dinámicas, requeriría necesariamente de un proceso prolongado de instrucción basado en los propios modelos dinámicos de evaluación y de mediación. En este sentido, ya hicimos referencia al mayor grado de correlación entre las puntuaciones dinámicas y las puntuaciones de rendimiento escolar, cuando el nivel de correspondencia entre las actividades contenidas en las pruebas dinámicas y las tareas escolares habituales era mayor (Sternberg y Grigorenko, 2003).

Estando básicamente de acuerdo con el contenido de estos análisis que ponen en duda la idoneidad del uso del rendimiento académico como criterio para evaluar la validez de predicción de un dispositivo de evaluación dinámica, y basando este acuerdo fundamentalmente en las diferencias sustanciales que pueden existir entre los procesos de intervención educativa puestos en práctica, así como entre los criterios de evaluación adoptados; nos hacemos eco igualmente de las palabras de Sternberg y Grigorenko cuando aluden a lo razonable que puede ser pensar que si el funcionamiento cognitivo de un sujeto se ha visto modificado en dirección positiva mediante un proceso de intervención, este hecho debería reflejarse de algún modo en su rendimiento escolar. Desde la perspectiva ecológica que hemos adoptado en nuestra propuesta, el rendimiento escolar, es decir, el funcionamiento cognitivo y socio-personal del sujeto en el contexto escolar, debería constituir sin lugar a dudas el criterio externo de validez del modelo dinámico propuesto. Claro está que la propia validez del criterio *rendimiento académico* puede ser muy distinta en función de cuáles sean los valores utilizados como criterio (calificaciones escolares, valoración del profesorado, puntuación en pruebas de aptitud, valoración del proceso de aprendizaje, etc.). Ya hicimos alusión anteriormente a la posibilidad de establecer un valor de predicción del *proceso de aplicación* de la prueba dinámica, con relación a la valoración *contextualizada* de un proceso de aprendizaje.

Hemos hecho referencia igualmente en el capítulo mencionado, así como en este mismo, a las premisas fundamentales sobre las que el análisis de la predicción de un modelo de evaluación dinámica, tendría un mayor sentido a nuestro juicio. En este sentido, la validación de un procedimiento de evaluación quedaría supeditada a la idoneidad del mismo para establecer dónde se encuentran las dificultades del sujeto y cómo es posible acceder a mejores cotas de funcionamiento. La predicción, valorada fundamentalmente con base en el proceso de aprendizaje, jugaría su papel precisamente en la determinación de patrones metodológicos de evaluación e intervención que demostraron ser eficaces en el proceso de aplicación dinámica.

Como ya quedó apuntado con anterioridad, en este trabajo hemos obtenido información del rendimiento escolar de los sujetos del GE a través de: (a) las calificaciones escolares en el área de Lengua al finalizar el curso; (b) la valoración cuantificable del profesorado acerca de una serie de habilidades relacionadas con los procesos contemplados en el dispositivo y con los criterios de evaluación del área de Lengua por parte de la Administración educativa; (c) el proceso de aprendizaje y de ajuste personal-social de los sujetos experimentales durante las sesiones de aplicación del dispositivo EDPL; (d) la puntuación obtenida en las pruebas postest, entendidas en este caso como pruebas de aptitud relacionadas con los contenidos escolares evaluados. Por otra parte, también hemos obtenido información del rendimiento académico de los sujetos del GC a través de los mismos procedimientos explicitados anteriormente, exceptuando lógicamente el que hace referencia al proceso de aplicación experimental del dispositivo.

De esta forma, se procedió a construir y a poner en circulación una plantilla (Anexo IX) que recogía, bien para cada grupo de aplicación experimental, o bien para cada grupo de sujetos del GC que pertenecían a un mismo Centro, los nombres de los sujetos que formaban parte de los mismos. El profesorado debía valorar, mediante el uso de una escala cualitativa en cuatro niveles, cada uno de los 7 criterios de evaluación propuestos. Estos criterios fueron formulados con base en los procesos implicados en la lectura contemplados en nuestra investigación, así como en los criterios de evaluación del área de Lengua relacionados especialmente con la comprensión lectora, que están contemplados en los Decretos de la Administración educativa. La valoración fue propuesta tanto a los profesores aplicadores del dispositivo como a los tutores de los cursos donde los sujetos se encontraban escolarizados. Igualmente, para los sujetos del GC, se solicitó la valoración de los tutores y del profesorado de apoyo a la integración que atendía al alumnado participante.

La gradación cualitativa propuesta para la valoración del profesorado fue la siguiente:

1. *Nivel bajo o muy bajo*
2. *Nivel medio-bajo*
3. *Nivel medio-alto*
4. *Nivel alto o muy alto*

Esta valoración del profesorado acerca del rendimiento escolar en el área de lengua debía realizarse en función del alumno, de su progreso individual, y se concretaba mediante el trazado de un círculo alrededor del número que mejor correspondiera al nivel de rendimiento del alumno. Así mismo, en la plantilla figuraban dos casillas junto a los números de la escala cualitativa, que hacían alusión al progreso que el docente había observado en el sujeto para cada uno de los criterios considerados. El profesor debía marcar con una cruz la casilla «P» si consideraba adecuado el progreso del alumno en el criterio de evaluación valorado, o «NP» si el sujeto no había progresado de forma adecuada.

Los criterios contemplados para esta valoración del rendimiento escolar en el área de Lengua fueron los siguientes:

- a) *Utiliza estrategias adecuadas en situaciones de diálogo en el aula: escuchar, respetar opiniones, expresar sus puntos de vista...*

b) Lee y comprende diferentes tipos de textos escolares adecuados a su nivel, destacando el tema y las ideas principales.

c) Utiliza estrategias de planificación en sus actividades de lectura, identificando la tarea que debe realizar y estableciendo unos objetivos.

d) Plantea dudas o preguntas al leer, se da cuenta de cuándo no comprende, formula conclusiones o comentarios de los textos leídos, evalúa su propia comprensión al terminar de leer.

e) Muestra una actitud positiva hacia la lectura.

f) Elabora esquemas o resúmenes de forma clara y ordenada, captando el sentido global del texto y las ideas principales.

g) Utiliza el Lenguaje escrito por propia iniciativa como medio para la adquisición de nuevos aprendizajes y como fuente de información.

Hemos obtenido la valoración del rendimiento escolar en el área de Lengua para 59 de los 60 sujetos que finalmente formaron parte del GE, así como un total de 64 valoraciones realizadas por los tutores y los profesores de apoyo del alumnado que formaba parte del GC. Igualmente hemos obtenido del profesorado colaborador las calificaciones finales en el área de Lengua de la mayoría de los sujetos participantes en el estudio.

Nos disponemos a continuación a exponer los resultados de los contrastes realizados para evaluar el valor de predicción diferencial del dispositivo EDPL con relación a las otras medidas obtenidas durante el desarrollo del estudio. Estas medidas serán, principalmente: (a) las puntuaciones pretest de las distintas pruebas criterio utilizadas; (b) las puntuaciones postest de las mismas pruebas; (c) las puntuaciones de mejora, obtenidas de la diferencia entre la puntuación postest y la pretest; (d) las valoraciones cuantificadas realizadas por los profesores tutores y de apoyo acerca del rendimiento escolar de los sujetos experimentales en el área de Lengua; (e) las valoraciones cuantificadas del progreso del alumnado participante en cada uno de los criterios evaluados en la plantilla antes mencionada; (f) las calificaciones escolares obtenidas en el área de Lengua por los sujetos; y, por último, (g) las propias puntuaciones dinámicas del dispositivo EDPL, obtenidas a raíz de la cuantificación de las valoraciones realizadas por los aplicadores durante el desarrollo de la aplicación experimental.

Un factor importante de la validez predictiva de una prueba dinámica vendría dado por la determinación de un mayor poder de predicción sobre un criterio como el rendimiento escolar, con relación a una prueba estática. En este sentido, hemos tratado de comparar el valor de predicción del dispositivo EDPL con el que aportarían las puntuaciones estáticas ofrecidas por las pruebas criterio. Hemos comenzado por la prueba de comprensión desarrollada en educación primaria, ECO 1. Se han considerado las siguientes puntuaciones como indicadoras del potencial dinámico del aprendizaje: (1) la puntuación media global obtenida en el dispositivo EDPL; y (2) la puntuación de mejora postest *menos* pretest. La puntuación postest, como medida post-instrucción, ha sido utilizada por diversos autores como medida del potencial de aprendizaje (Resing, 2000); aunque tendremos en cuenta esta medida a la hora de establecer las comparaciones, sobre todo en lo relativo a la puntuación inicial, no debemos olvidar que

se trata de una medida estática. Las dos medidas dinámicas contempladas deberían aportar un significado adicional a la predicción y la explicación del rendimiento escolar de los sujetos participantes en el estudio con relación a la información que puede obtenerse de las pruebas iniciales.

Valor predictivo en la prueba de comprensión ECO 1

En las siguientes tablas exponemos los datos descriptivos y los resultados de los contrastes para el grupo experimental en educación primaria.

Tabla 76. Valores y puntuaciones obtenidas por el GE en las diversas medidas contempladas

Medidas observadas	Valores medios	Grupo experimental (N)
P. dinámica EDPL	95,085	32
Pretest ECO 1	84,184	32
Postest ECO 1	85,796	30
Puntuación de mejora Post – Pret	0,598	30
Valoración tutor (sobre 4)	1,830	32
Suma valoraciones progreso (sobre 7)	2,810	32
Valoración apoyo (sobre 4)	2,084	22
Suma valoraciones progreso (sobre 7)	4,95	22
Calificaciones Lengua (sobre 10)	4,42	19

Tabla 77. Correlaciones entre los valores y puntuaciones obtenidos por el GE en las diversas medidas contempladas.

Los tamaños de las muestras de correlación están en función de los recogidos en la tabla anterior

	Dinám. EDPL	Pretest ECO 1	Postest ECO 1	Mejora	Valora tutor	∑ valora progreso	Valora apoyo	∑ valora progreso	Notas Lengua
P. dinámica EDPL		-0,503**	-0,198	0,228	0,616**	0,391*	0,443*	0,282	-0,761**
Pretest ECO1	-0,503**		0,431*	-0,466**	-0,201	0,178	-0,127	0,031	0,449
Postest ECO1	-0,198	0,431*		0,597**	0,120	-0,149	0,392	0,280	-0,060
P. Mejora Post – Pret	0,228	-0,466**	0,597**		0,320	-0,266	0,521*	0,438*	-0,376
Valoración tutor (1-4)	0,616**	-0,201	0,120	0,320		0,641**	0,361	0,180	-0,312
∑ valoración progreso	0,391*	0,178	-0,149	-0,266	0,641**		0,179	0,161	0,109
Valoración apoyo (1-4)	0,443*	-0,127	0,392	0,521*	0,361	0,179		0,591**	-0,542*
∑ valoración progreso	0,282	0,031	0,280	0,438*	0,180	0,161	0,591**		-0,111
Notas Lengua	-0,761**	0,449	-0,060	-0,376	-0,312	0,109	-0,542*	-0,111	

* La correlación es significativa al nivel de 0,05; ** La correlación es significativa al nivel de 0,01

La inspección minuciosa de los datos recogidos en la tabla anterior nos ofrece algunos datos relevantes que pueden ayudarnos a comprender mejor y a matizar algunos de los resultados obtenidos en nuestro estudio, así como a establecer las primeras conclusiones sobre el valor predictivo del dispositivo de evaluación dinámica aplicado.

En primer lugar, debemos hacer notar la alta correlación, muy significativa estadísticamente, que presenta la puntuación dinámica EDPL con la valoración del tutor sobre el rendimiento escolar de los sujetos. Esta significación se da igualmente con el valor extraído de la suma de valoraciones positivas sobre el progreso de los alumnos, así como sobre la valoración del profesorado de apoyo sobre el rendimiento. En cambio, es igualmente notorio el valor negativo de las correlaciones con las puntuaciones pretest (muy significativa) y postest, así como con las calificaciones escolares (también muy significativa y negativa). Esta divergencia entre las correlaciones de la puntuación dinámica con el pretest y el postest, se habría visto reflejada sin embargo en el valor de correlación de la puntuación de EDPL con la mejora. Como podemos observar, esta correlación pasa a ser positiva, con un valor de $r = 0,228$; valor que concuerda con la idea de que las puntuaciones dinámicas, como expresión de la valoración de un proceso de aprendizaje, guardarían poca relación con puntuaciones que expresan la valoración de un producto y, en cambio, guardarían una mayor afinidad con puntuaciones que, de alguna forma, recojan la valoración de un proceso de cambio. En este sentido, los resultados expresan de forma clara una mayor sintonía entre la puntuación dinámica y la puntuación de mejora, sintonía que no se da en cambio con las puntuaciones obtenidas de la pasación de las pruebas estáticas.

Con relación a las calificaciones escolares, podemos observar que entre las valoraciones efectuadas por el propio profesorado de los alumnos acerca de su rendimiento escolar en Lengua, y las calificaciones académicas para esta misma materia, existe una correlación negativa. El valor de esta correlación es bastante alto en el caso del profesorado tutor ($r = -0,312$) y llega a ser significativa en el caso del profesorado de apoyo ($r = -0,542$). Hay que hacer notar igualmente la correlación negativa de las calificaciones escolares con la puntuación dinámica; el valor de esta correlación negativa llega a ser muy significativa. De igual forma, es también negativa con relación a la puntuación de mejora, la puntuación postest y la valoración del progreso efectuada por el profesorado de apoyo. Sin embargo, podemos comprobar que existe una alta correlación positiva de las calificaciones con la puntuación pretest. Posteriormente, en el capítulo de Discusión, abordaremos las posibles explicaciones que podrían estar detrás de estos resultados.

En cuanto a los valores de correlación de la puntuación de mejora, podemos observar en la tabla que ésta mantiene correlación positiva con la puntuación dinámica, con la puntuación postest, llegando a ser muy significativa, con la valoración del tutor sobre el rendimiento escolar, y con la valoración del profesorado de apoyo sobre el rendimiento y sobre el progreso de los sujetos, ambas también significativas. En cambio, la correlación es negativa y significativa con el pretest, con la valoración del progreso expresada por los tutores y con las calificaciones escolares. El análisis de estos datos nos ofrece información cualitativa muy valiosa, que analizaremos también con mayor profundidad cuando abordemos el análisis y la discusión de los resultados.

Como hemos apuntado, la correlación del pretest con el postest es significativa. En este sentido, no nos parece extraño que se dé esta correlación. Aunque puede variar sustancialmente en función del tipo de pruebas que se apliquen, la existencia de

correlación entre las puntuaciones obtenidas por los mismos sujetos en la misma prueba, puede ser de hecho natural y esperable. Es más significativo para nosotros que el pretest correlacione negativamente con la puntuación de mejora y, en cambio, que el postest correlacione positiva y muy significativamente con esta misma puntuación. Otro dato relevante con respecto a la puntuación inicial lo constituye la correlación negativa que mantiene con las valoraciones efectuadas por los tutores y por el profesorado de apoyo. Sin embargo, la correlación de la puntuación postest con estas mismas valoraciones pasa a ser positiva. Este dato no nos debe extrañar, sobre todo si tenemos en cuenta que las valoraciones fueron realizadas al finalizar el curso, justo cuando se desarrollaron las pruebas postest. Lógicamente, las valoraciones del profesorado sobre el rendimiento escolar en cada uno de los criterios contemplados y sobre el progreso de los sujetos en estos criterios, debían estar más en consonancia con la actuación en el postest. El análisis de los datos del grupo control nos revela resultados similares en este contraste; el valor de la correlación entre el rendimiento escolar y el pretest ($N = 27$) es negativo ($r = -0,129$), mientras que es positivo y significativo con el postest ($r = 0,453$). Ocurre igual con la correlación entre la suma de las valoraciones positivas sobre el progreso del alumnado y las puntuaciones pretest y postest, es decir, correlación negativa —que llega a ser significativa— para el pretest, y correlación positiva y muy significativa para el postest.

Para el GC tomado en su totalidad, la valoración promedio del tutor sobre el rendimiento escolar ($N = 27$) fue de 1,667 y el valor medio de la suma sobre las valoraciones positivas de progreso fue de 2,33; así mismo, la valoración media del profesorado de apoyo ($N = 7$) fue de 1,735 y la valoración media del progreso fue de 4,00. El promedio de las calificaciones escolares ($N = 22$) arrojó un resultado de 4,23. Paradójicamente, a pesar de que este grupo obtuvo ganancias postest, los resultados de todas las valoraciones efectuadas por el profesorado sobre el rendimiento escolar y sobre el progreso del alumnado, estuvieron por debajo de los que se obtuvieron con los sujetos del GE. Hemos efectuado contrastes de medias entre las valoraciones otorgadas al GE y al GC mediante la prueba *t* para muestras independientes; en este sentido, a pesar de las diferencias observadas, los resultados de estos contrastes no revelan la existencia de diferencias significativas estadísticamente. Otro dato relevante —que será discutido con posterioridad— con relación al análisis de las correlaciones entre el rendimiento escolar y las puntuaciones obtenidas por el GC, es el que hace referencia a las calificaciones escolares. En este caso, las notas en Lengua correlacionan negativamente con el pretest, al contrario de lo que ocurriera en el GE, y lo hacen positiva y significativamente con el postest ($N = 22$). También correlacionan de forma significativa con la puntuación de mejora y muy significativamente con la valoración del rendimiento efectuada por el tutor, así como con la valoración del progreso ($r = 0,589$ y $r = 0,587$ respectivamente).

Análisis de regresión en la prueba ECO 1

Hemos efectuado diversos análisis de regresión con el fin de complementar la información sobre el valor predictivo del dispositivo EDPL. Estos análisis los hemos realizado sobre: (a) el rendimiento escolar en Lengua, medido con la valoración del tutor; (b) el rendimiento escolar en Lengua, medido con la valoración del profesorado de apoyo; (c) el rendimiento escolar medido con las calificaciones finales en el área de Lengua; y (d) la puntuación postest. Por otra parte, hemos efectuado igualmente unos análisis de regresión en función del grupo de alumnos del GE que obtuvo mejoras en el

postest (mejoradores), así como de aquellos que no consiguieron ganancias (no mejoradores).

Con relación al rendimiento académico, medido con la valoración del tutor, la puntuación dinámica EDPL predijo este valor con una R^2 de 0,379. El contraste para valorar el ajuste del modelo arrojó un valor F de 18,323 y una significación de 0,000. Por su parte, como hemos anotado anteriormente, la puntuación pretest correlacionó negativamente con el rendimiento escolar; la puntuación postest explicaba por sí sola el 0,14% de la varianza observada, mientras que la puntuación de mejora también predecía tan sólo un 1,02% de la varianza. Los resultados del análisis de regresión por pasos muestran que tan sólo la puntuación dinámica entra en la ecuación, con un valor R^2 de 0,389; el resto de variables quedaron excluidas al no sobrepasar el criterio del valor de probabilidad 0,05 para la prueba F. De esta forma, mientras que para el GC la puntuación postest y la puntuación de mejora predecían de forma significativa el rendimiento escolar, para el GE fue la puntuación dinámica obtenida de la aplicación del dispositivo, el único valor que mantuvo una relación significativa con el rendimiento.

El análisis de regresión efectuado con relación a la valoración sobre el progreso de los sujetos que realizaron los tutores, mostró la inclusión de dos variables en la ecuación. En primer lugar, la puntuación dinámica, con un valor de correlación de $r=0,433$ y una R^2 de 0,188; por sí sola explicaría de forma significativa la varianza observada, con un valor F de 6,464 y una probabilidad asociada de 0,017. Cuando se introduce en la ecuación la puntuación pretest, que inicialmente tenía un valor de correlación de $r=0,134$, se produce un incremento en R^2 de 0,150, explicando conjuntamente ambas variables el 33,8% del progreso valorado en los sujetos. El ajuste del modelo cuando son introducidas las dos variables es muy significativo ($F = 6,866$; $p < 0,004$). Paradójicamente, la puntuación de mejora y la puntuación postest correlacionaron negativamente con la valoración del progreso.

En cuanto al rendimiento académico valorado por el profesorado de apoyo, la puntuación dinámica volvía a actuar como predictora, explicando por sí sola un 19,6% de la varianza. El valor F fue de 4,883 y la significación asociada de 0,039. Por su parte, la puntuación pretest correlacionaba negativamente con la valoración del rendimiento escolar efectuada por los profesores de apoyo. La puntuación postest predecía el rendimiento con un valor R^2 de 0,154, siendo el valor F de 3,453 y la probabilidad asociada de 0,079. Por último, la puntuación de mejora predecía un 27,2% de la varianza observada, con un valor F de 7,089 y una significación de 0,015. Cuando se introdujeron todas las variables en un análisis de regresión por pasos, sólo la puntuación de mejora salvó el criterio de entrada, con un valor R^2 de 0,272.

La valoración sobre el progreso del alumnado efectuada por el profesorado aplicador fue sometida igualmente a un análisis de regresión por pasos. Sólo la puntuación dinámica entró en la ecuación, con un valor de correlación de $r=0,464$ y una R^2 de 0,215. El contraste realizado para evaluar el ajuste del modelo fue significativo, con un valor F de 5,208 y una probabilidad asociada de 0,034. Se excluyeron de la ecuación la puntuación pretest, con un valor de correlación negativa ($r=-0,142$); la puntuación postest ($r=0,280$); y la puntuación de mejora, con un valor de correlación bastante alto ($r=0,438$). Cuando se introdujeron todas las variables en bloque, la puntuación postest quedó excluida nuevamente de la ecuación, explicando el conjunto de variables

incluidas el 32,5% de la varianza observada; sin embargo, este contraste no resultó significativo estadísticamente ($F = 2,725$; $p < 0,076$).

El rendimiento escolar valorado con las calificaciones obtenidas en el área de Lengua fue explicado en un 20,1% por las puntuaciones pretest, con un valor para la prueba F de 4,281 y una significación asociada de 0,054. Las puntuaciones posttest predecían tan sólo un 0,04% de la varianza observada, mientras que la puntuación dinámica y la puntuación de mejora correlacionaban negativamente con las notas finales. Cuando se realizó el análisis de regresión por pasos, ninguna variable fue introducida en la ecuación al no sobrepasar el criterio de probabilidad de la prueba F. Estos resultados vienen a confirmar con más rotundidad si cabe lo apuntado anteriormente, en el sentido de que el rendimiento escolar basado en las calificaciones escolares finales no ha funcionado adecuadamente como criterio para establecer el valor predictivo del tratamiento.

En cuanto a la puntuación posttest, los resultados del análisis de la regresión cuando las variables fueron introducidas individualmente, mostraron un valor R^2 de 0,356 para la puntuación de mejora; resultando significativa la proporción de varianza explicada por esta puntuación. El valor de contraste de la prueba F fue de 15,511, con una significación asociada de 0,000. Por su parte, el análisis efectuado para el pretest, arrojó un valor de R^2 de 0,186. El modelo es significativo, con un valor F de 6,389 y una significación de 0,017. Cuando se procedió a realizar el análisis de regresión por pasos, los resultados mostraron que la puntuación de mejora aparecía en primer lugar, con el valor R^2 mencionado más arriba, seguida de la puntuación pretest. El incremento en R^2 resultó ser muy significativo (0,644), explicando ambas variables conjuntamente el 100% de la varianza observada en el posttest.

Por su parte, la determinación del conjunto de puntuaciones que mejor predice la mejora en el posttest, nos podría otorgar una valoración de la posible idoneidad de las puntuaciones dinámicas sobre las estáticas. En este sentido, la correlación de las puntuaciones de mejora con las puntuaciones pretest, posttest y con las puntuaciones dinámicas, debía darnos esta información. Los resultados mostraron que la puntuación pretest, así como las calificaciones escolares, correlacionaban negativamente con la mejora. Por otra parte, la puntuación posttest predecía de forma individual la mejora con un valor R^2 de 0,356, un valor para la prueba F de 15,511 y una significación asociada de 0,000. La puntuación dinámica explicaba la de mejora con un valor R^2 de 0,052, un valor F de 1,541 y una significación de 0,225. A pesar de estos valores pequeños para la puntuación EDPL, los resultados del análisis de regresión por pasos nos mostraron que las puntuaciones posttest y las puntuaciones dinámicas EDPL explicaban conjuntamente el 48,2% de la varianza observada, produciéndose un incremento en R^2 de 0,125 cuando se introducía la puntuación dinámica en la ecuación. En cualquier caso, el establecimiento de la comparación lógica entre las puntuaciones estáticas del pretest y las puntuaciones dinámicas, que corresponden a momentos anteriores a la determinación de las puntuaciones de mejora, es claramente favorable a las puntuaciones dinámicas.

El análisis de regresión para los grupos de mejoradores y no mejoradores

Como apuntábamos anteriormente, hemos realizado los cálculos del análisis de regresión en función del grupo de sujetos de la muestra experimental que obtuvo

mejoras postest en la prueba analizada, en este caso, la prueba de comprensión ECO 1. El grupo de alumnos que, siguiendo la terminología de Budoff, habíamos denominado *mejoradores*, y que estaba compuesto por 8 sujetos, obtuvo en el pretest una puntuación de 76,544; mientras que en el postest fue de 85,796. Su puntuación dinámica EDPL fue de 105,515; la puntuación de mejora fue de 9,252; y, por último, la valoración media de los tutores fue de 2,25.

La introducción individual de las variables con relación al rendimiento académico, medido con la valoración del tutor, otorgó los siguientes resultados: la puntuación dinámica correlacionó con un valor de $r.0,553$; la R^2 fue de $0,306$ y el valor de la prueba F para el contraste fue de $2,646$, con una significación asociada de $0,155$. La puntuación pretest obtuvo un valor medio de correlación positiva ($r.0,359$; sig. $0,191$); mientras que la postest correlacionó positiva pero débilmente con el rendimiento ($r.0,091$; sig. $0,415$). Por su parte, la puntuación de mejora correlacionó con un valor negativo ($r.-0,183$). Cuando los cálculos fueron realizados mediante un análisis de regresión por pasos, los resultados mostraron que ninguna de ellas sobrepasaba el criterio ($p < 0,05$) para ser introducida en la ecuación. A pesar de esto, para este subgrupo de sujetos mejoradores, pudimos constatar que las puntuaciones dinámicas explicaban y predecían mejor que ninguna otra puntuación, la valoración del rendimiento efectuado por los tutores. Por su parte, el subgrupo de los no mejoradores, compuesto por 3 sujetos, obtuvo las siguientes valoraciones: su puntuación pretest fue de $80,937$, mientras que la postest fue $78,694$; la puntuación dinámica EDPL fue de $97,492$ y la puntuación de mejora (diferencia postest-pretest) fue de $-2,243$. La valoración media del rendimiento obtuvo un valor de $1,571$. En este último dato aparece como reveladora la menor valoración media que recibieron los sujetos no mejoradores con respecto a los mejoradores. Por otra parte, los resultados del análisis de regresión por pasos mostraron para este grupo que ninguna variable fue introducida en la ecuación, obteniendo la puntuación dinámica el valor de correlación más alto ($r.0,615$), con una R^2 de $0,378$ ($F = 0,608$; $p < 0,578$). Por otra parte, tanto la puntuación postest ($r.-0,665$), como la puntuación de mejora ($r.-0,189$), obtuvieron valores de correlación negativa con el rendimiento escolar. Por último, la puntuación pretest también correlacionó negativamente con el rendimiento para este subgrupo ($r.-0,327$), al contrario de cómo lo hiciera para el subgrupo de mejoradores. En definitiva, la puntuación dinámica actuó como el valor que mejor predecía el rendimiento escolar, tanto en el grupo de mejoradores como en el de no mejoradores; sin embargo, estos valores de predicción no llegaron a ser significativos estadísticamente en ninguno de estos grupos, probablemente debido a la escasa potencia de la muestra observada. Cuando realizamos los cálculos para el grupo de sujetos que obtuvieron pérdidas postest, independientemente de su puntuación inicial ($N = 16$), los resultados del análisis de regresión por pasos mostraron para este grupo ampliado que la puntuación dinámica, con un valor de correlación de $r.0,786$, predecía en un $61,7\%$ el rendimiento escolar, siendo la única variable introducida en la ecuación; el valor de ajuste del modelo fue de $22,588$, con una significación asociada de $p < 0,000$. La valoración de los tutores parece haber ido, de esta forma, en línea con las valoraciones efectuadas por los aplicadores sobre el proceso de aprendizaje realizado durante la implementación del dispositivo.

En cuanto al rendimiento escolar valorado por el profesorado de apoyo, el grupo de mejoradores, compuesto en este caso por 5 sujetos, obtuvo una valoración media de $2,514$; una puntuación pretest de $76,152$ y una postest de $85,572$; su puntuación dinámica fue de $101,442$ y la de mejora de $9,420$. Analizadas de forma individual, los

valores de correlación fueron exiguos o negativos ($r=0,126$ para el pretest; $r=-0,006$ para el postest; $r=-0,070$ para la puntuación de mejora; y $r=-0,439$ para la puntuación dinámica). Ninguno de estos valores se acercaba a la significatividad estadística; el valor R^2 mayor fue de $0,192$ para la puntuación dinámica, que paradójicamente obtuvo un valor de correlación negativo con respecto al rendimiento valorado por el profesorado de apoyo. Cuando se efectuó un análisis de regresión por pasos, obviamente no fue introducida ninguna de las variables en la ecuación, ya que ningún contraste se acercaba al valor de probabilidad criterio. Por su parte, el grupo de no mejoradores, compuesto por 2 sujetos, obtuvo una valoración media de $1,786$; una puntuación pretest de $81,310$ y una postest de $79,068$; su puntuación dinámica fue de $92,583$ y la puntuación de mejora (diferencia postest-pretest) de $-2,243$. Hemos vuelto a realizar los cálculos para todos los sujetos que obtuvieron pérdidas, aunque no obtuvieran puntuaciones inicialmente bajas. En este caso, el subgrupo se eleva a 10 sujetos, que obtuvieron una valoración media de $1,729$; una puntuación pretest de $87,142$ y una postest de $81,535$; su puntuación dinámica fue de $81,928$ y la puntuación de mejora (diferencia postest-pretest) de $-5,607$. Al igual que ocurriera con la valoración de los tutores, podemos observar que la valoración media del rendimiento escolar de los no mejoradores desciende notablemente con respecto a la de los mejoradores. En el caso de los no mejoradores, la puntuación dinámica ($r=0,667$), cuando es introducida individualmente, explica el $44,4\%$ del rendimiento escolar; la puntuación de mejora ($r=0,236$) obtiene un valor R^2 de $0,056$, y la puntuación postest ($r=0,144$) una R^2 de $0,021$. Por su parte, el pretest correlacionaba negativamente ($r=-0,078$) con el rendimiento valorado. El análisis de regresión por pasos muestra que tan sólo la puntuación dinámica explica significativamente el rendimiento escolar, obteniendo los valores de correlación y R^2 indicados anteriormente; el ajuste del modelo se expresa con un valor F de $6,398$ y una probabilidad asociada de $0,035$. Ninguna otra variable entra en la ecuación.

Con relación a las calificaciones escolares, el grupo de mejoradores, integrado también por 5 sujetos, obtuvo las siguientes valoraciones: $4,20$ puntos de calificación media; una puntuación pretest de $76,152$ y una postest de $85,572$; una puntuación dinámica de $101,442$ y una puntuación de mejora de $9,420$. El análisis de regresión para cada una de las variables contempladas mostró que la puntuación dinámica correlacionó negativamente ($r=-0,805$) con las notas, con un valor R^2 de $0,649$ y un valor F de $5,535$ ($p<0,100$). La puntuación de mejora también correlacionó negativamente con las notas ($r=-0,523$), con un valor R^2 de $0,274$. En cambio, el pretest correlacionaba positivamente ($r=0,875$) y predecía las calificaciones con un valor R^2 de $0,766$. El valor del contraste se encontraba en el límite de la significatividad estadística, obteniendo un valor F de $9,800$ y una probabilidad asociada de $0,052$. Por su parte, el postest también correlacionó negativamente con las calificaciones ($r=-0,108$), obteniendo un valor R^2 de $0,012$. Cuando se efectuó el análisis de regresión por pasos, ninguna de las variables fue introducida en la ecuación al no sobrepasar el criterio de entrada. En síntesis, las notas del grupo de mejoradores sólo correlacionaron positivamente con las puntuaciones estáticas del pretest, obtenidas en la prueba de comprensión realizada 6 meses antes de las calificaciones escolares; sin embargo, aquellas puntuaciones dinámicas, que de uno u otro modo reflejaban el proceso de aprendizaje y las mejoras obtenidas, no se vieron reflejadas en las calificaciones finales.

Por otra parte, para los no mejoradores ($N = 2$), la calificación media se situó en $4,50$ puntos. Ya hemos hecho mención del resto de puntuaciones obtenidas por este reducido

número de los no mejoradores. Al igual que en ocasiones anteriores, esta muestra nos ha permitido realizar cálculos conducentes a establecer relaciones con una mínima fiabilidad entre las distintas valoraciones. Por su parte, para el grupo ampliado de no mejoradores ($N = 9$), la calificación media fue de 4,56; la puntuación pretest de 86,668 y la del postest de 80,438; la puntuación dinámica obtuvo 83,395 y la de mejora obtuvo un valor negativo de -6,230. El análisis de regresión para cada variable mostró que sólo la puntuación pretest correlacionó positivamente con las notas de este grupo, aunque de forma no significativa ($r.0,182$), con un valor R^2 de 0,033. En cambio, las puntuaciones dinámicas, del postest y de mejora, correlacionaron negativamente con las notas. Para las primeras, con un valor de correlación $r.-0,713$ y un valor R^2 de 0,508, el contraste con la prueba F resultaba significativo ($F = 7,237$; sig. 0,031). El análisis de regresión por pasos no mostró obviamente ninguna que explicara de forma significativa las calificaciones otorgadas a este grupo. Un aspecto interesante de la comparación entre los grupos de mejoradores y no mejoradores, fue el contraste entre las valoraciones otorgadas por los tutores y las calificaciones. Ya habíamos comentado que la correlación para el GE completo era sorprendentemente negativa. Cuando hemos realizado el contraste para estos dos grupos por separado, hemos constatado que para el grupo de mejoradores, la correlación de las notas con la valoración media del tutor es positiva ($r.0,129$), mientras que es negativa y significativa para el grupo de no mejoradores ampliado ($r.-0,693$). La valoración de los tutores para el primer grupo (2,171) también superó ampliamente a la valoración otorgada a los sujetos que no mejoraron (1,269). Estos resultados matizan en algo el dato de correlación negativa entre valoración del tutor y calificaciones, derivando el mismo hacia el grupo de no mejoradores.

Por último, con relación a la puntuación postest, los resultados del análisis de regresión para cada variable mostraron que el postest era predicho por el pretest en un 35,5% en lo que respecta al grupo de mejoradores ($N = 8$). El valor de la prueba F fue de 3,297 y la significación asociada de 0,119. La puntuación de mejora correlacionó 0,862 con el postest y obtuvo un valor R^2 de 0,742 ($F = 17,293$; $p < 0,006$). Por su parte, la correlación entre la puntuación dinámica y el postest fue positiva ($r.0,313$), con un valor R^2 de 0,098. El análisis de regresión por pasos mostró que la conjunción de la puntuación pretest y la puntuación de mejora, al igual que ocurriera con el grupo experimental completo, explicaba la totalidad de la varianza observada. El modelo sufría un incremento en R^2 de 0,258 cuando era introducida la puntuación pretest, con lo que unido al valor R^2 de la mejora, predecía la totalidad de la puntuación postest. Para el grupo de no mejoradores, con sus 3 integrantes, la puntuación dinámica correlacionó negativamente con el postest ($r.-0,999$); la puntuación pretest obtuvo un valor de correlación también negativo de $r.-0,500$, y la puntuación de mejora correlacionó positiva y significativamente con el postest ($r.0,866$), obteniendo un valor R^2 de 0,750, aunque no se produjera finalmente un ajuste significativo del contraste del modelo, con un valor F de 3,000 y una significación asociada de 0,333. Por su parte, para el grupo ampliado de sujetos no mejoradores ($N = 16$), la puntuación dinámica correlacionó también negativamente con el postest ($r.-0,026$), mientras que el pretest ($r.0,538$) predecía en un 29% la puntuación final, con un valor de prueba F de 5,713 y una probabilidad asociada de 0,031. Por su parte, la puntuación de mejora correlacionó positiva y significativamente con el postest ($r.0,671$), con un valor R^2 de 0,451 y una F de 11,481 ($p < 0,004$). El análisis de regresión por pasos revela nuevamente que la conjunción de ambas puntuaciones, mejora y pretest, explicarían la totalidad de la varianza observada. La primera variable introducida es la puntuación de mejora,

produciéndose un incremento en R^2 de 0,549 cuando entra en la ecuación la puntuación pretest. El modelo, por tanto, explica de forma significativa la puntuación posttest.

Valor predictivo en la prueba de comprensión ECOS

Exponemos a continuación los datos descriptivos y los resultados de los contrastes para el grupo experimental en educación secundaria.

Tabla 78. Valores y puntuaciones obtenidas por el GE en las diversas medidas contempladas

Medidas observadas	Valores medios	Grupo experimental (N)
P. dinámica EDPL	105,874	28
Pretest ECOS	88,251	28
Posttest ECOS	95,375	25
Puntuación de mejora Post – Pret	7,928	25
Valoración tutor (sobre 4)	2,402	27
Suma valoraciones progreso (sobre 7)	4,78	27
Valoración apoyo (sobre 4)	2,316	19
Suma valoraciones progreso (sobre 7)	4,26	19
Calificaciones Lengua (sobre 10)	5,74	27

Tabla 79. Correlaciones entre los valores y puntuaciones obtenidos por el GE en las diversas medidas contempladas.

Los tamaños de las muestras de correlación están en función de los recogidos en la tabla anterior

	Dinám. EDPL	Pretest ECOS	Posttest ECOS	Mejora	Valora tutor	∑ valora progreso	Valora apoyo	∑ valora progreso	Notas Lengua
P. dinámica EDPL		-0,220	-0,159	-0,048	0,350	0,388*	0,489*	0,630**	0,242
Pretest ECOS	-0,220		0,791**	-0,018	0,308	0,246	0,055	-0,303	0,253
Posttest ECOS	-0,159	0,791**		0,598**	0,302	0,171	0,091	-0,138	0,123
P. Mejora Post – Pret	-0,048	-0,018	0,598**		0,097	0,016	0,047	-0,005	-0,168
Valoración tutor (1-4)	0,350	0,308	0,302	0,097		0,764**	1,000**	0,741**	0,683**
∑ valoración progreso	0,388*	0,246	0,171	0,016	0,764**		0,741**	1,000**	0,686**
Valoración apoyo (1-4)	0,489*	0,055	0,091	0,047	1,000**	0,741**		0,741**	0,685**
∑ valoración Progreso	0,630**	-0,303	-0,138	-0,005	0,741**	1,000**	0,741**		0,659**
Notas Lengua	0,242	0,253	0,123	-0,168	0,683**	0,686**	0,685**	0,659**	

* La correlación es significativa al nivel de 0,05; ** La correlación es significativa al nivel de 0,01

La inspección de los datos que hemos analizado muestra de entrada algunas diferencias importantes con respecto a los que hemos obtenido en educación primaria. Iremos paulatinamente profundizando en estas diferencias a medida que nos adentramos en la descripción y el análisis de los contrastes efectuados.

En primer lugar hemos de mencionar la correlación positiva que mantiene la puntuación dinámica con todas las valoraciones sobre el rendimiento escolar efectuadas por el profesorado. Para tres de las cuatro efectuadas, esta correlación es significativa o muy significativa estadísticamente, destacando los contrastes con la valoración sobre el rendimiento y sobre el progreso efectuada por el profesorado aplicador. Por otra parte, al igual que ocurriera en primaria, la puntuación dinámica correlacionó negativamente con el pretest y con el postest, aunque en este caso, los valores no llegan a ser tan altos. También correlacionó negativamente con la puntuación de mejora, lo que constituye un dato un tanto sorprendente, que será objeto de análisis posteriormente. Podemos constatar, así mismo, que las valoraciones efectuadas por el profesorado aplicador sobre la actuación de los sujetos en las sesiones del dispositivo EDPL —reflejadas a nivel cuantitativo en las puntuaciones dinámicas—, han estado muy por encima de los valores obtenidos por los mismos sujetos en las pruebas de criterio; lo que unido al sentido claramente positivo de la correlación entre la puntuación dinámica y las valoraciones del rendimiento efectuadas por los docentes, apunta decididamente a que las pruebas estáticas no habrían logrado recoger parte de la información relevante relacionada con los procesos evaluados. Por último, con respecto a la puntuación dinámica, es destacable que, al contrario de lo que ocurriera en primaria, la correlación con las calificaciones finales haya sido positiva ($r=0,242$). En este caso, las notas sí muestran niveles de correlación positivos y muy significativos con la valoración del tutor, lo que era inicialmente lógico y esperable. Siguiendo con el criterio de rendimiento escolar basado en las calificaciones, hemos de destacar las correlaciones muy significativas que mantienen con todas las valoraciones efectuadas por el profesorado, tanto por los tutores como por el profesorado de apoyo. Por otra parte, las calificaciones correlacionan positivamente también con las puntuaciones pretest y postest, aunque los valores no son muy altos ($r=0,253$ y $r=0,123$, respectivamente). La puntuación de mejora es la única variable con la que mantiene una correlación negativa.

Otro dato relevante es la alta correlación que mantiene la puntuación pretest con la postest ($r=0,791$); sin embargo, con relación a la puntuación de mejora y la puntuación dinámica, la correlación es negativa. Este dato será discutido posteriormente, ya que desde un enfoque dinámico, el hecho de que las puntuaciones pretest guarden o no relación positiva con las puntuaciones dinámicas, podría informar acerca de la validez del tratamiento aplicado en función de los cambios que hubiera logrado provocar al margen de cuál fuera el punto de partida del sujeto. También expresa bajas correlaciones o correlaciones negativas con las valoraciones sobre el rendimiento escolar efectuadas por el profesorado aplicador del dispositivo; sin embargo, con las valoraciones de los tutores las correlaciones son moderadas ($r=0,308$ para la valoración sobre el rendimiento y $r=0,246$ para la valoración sobre el progreso). Si para el pretest, la única correlación significativa estadísticamente era la que mantenía con la puntuación postest, para ésta última, la única correlación significativa se produce con las puntuaciones de mejora. Con las valoraciones del tutor, el postest mantiene una correlación moderada ($r=0,302$), siendo el resto correlaciones de nivel bajo o incluso negativo.

Por otra parte, los datos observados avalan notablemente la validez del instrumento utilizado para realizar las valoraciones; por una parte, está la ya comentada correlación positiva y muy significativa que guardan todas las valoraciones del profesorado con las calificaciones; por otra, las correlaciones entre las valoraciones sobre el rendimiento y sobre el progreso de los alumnos, que también han sido muy significativas, tanto para el tutor como para el profesorado de apoyo. Este dato hablaría de la coherencia en las valoraciones efectuadas por el profesorado, al tiempo que sobre la validez del instrumento. Por último, tanto la correlación entre las valoraciones del tutor y las del profesor de apoyo sobre el rendimiento, como la correlación entre las valoraciones de ambos profesionales sobre el progreso de los sujetos, ha tenido el máximo valor, $r.1,000$.

Por lo que respecta al GC tomado al completo, la valoración promedio del tutor sobre el rendimiento escolar ($N = 27$) fue de 2,085 y el promedio de la suma de las valoraciones positivas de progreso fue de 2,59; en cuanto a las valoraciones del profesorado de apoyo, éstas no pudieron ser recogidas para este grupo, por lo que sólo contamos con las valoraciones efectuadas por los tutores. Por otra parte, el promedio de las calificaciones escolares ($N = 22$) fue de 4,64. En cuanto a las correlaciones, es destacable que el rendimiento escolar valorado por el tutor correlaciona de forma muy significativa con la puntuación de mejora ($r.0,597$) y con el postest ($r.0,672$). También correlaciona positivamente con el pretest ($r.0,304$), aunque de forma no significativa. Al igual que ocurriera con el grupo de primaria, parece ser que para el GC en esta etapa, la puntuación postest y la de mejora podrían explicar la mayor parte de la varianza observada. La consideración, en cambio, de la puntuación dinámica en el GE contribuye de forma significativa a explicar el rendimiento escolar. Por otra parte, en el GC el pretest también correlaciona muy significativamente con el postest ($r.0,747$), y lo hace positivamente con la mejora ($r.0,121$), aunque el valor de la correlación sea bajo. Con relación a las calificaciones, éstas correlacionan negativamente con el pretest ($r.-0,072$) y positivamente, aunque con coeficientes bajos y moderados, con el resto de valores y puntuaciones: con el postest ($r.0,137$), con la puntuación de mejora ($r.0,314$), con el rendimiento valorado por el tutor ($r.0,375$), y, por último, con la valoración sobre el progreso ($r.0,223$).

Otro aspecto destacable además de las correlaciones, es el que hace referencia a la diferencia entre las valoraciones efectuadas por el profesorado sobre el rendimiento escolar y el progreso del alumnado en los grupos experimental y control, así como en cuanto a las calificaciones escolares. El hallazgo de diferencias importantes entre las valoraciones de los dos grupos contribuiría a informar sobre el rendimiento de los sujetos más allá de su actuación en las pruebas criterio. En este sentido, en el GE anotábamos anteriormente que el profesorado tutor había otorgado 2,402 puntos de promedio, puntuación sensiblemente superior a los 2,085 puntos promedio que recibieron los sujetos en el GC. Hemos contrastado esta diferencia de medias mediante la prueba *t* para muestras independientes; los resultados muestran un valor *t* de 1,576, con una significación asociada de 0,121 para 52 grados de libertad. Por otra parte, la valoración promedio de los tutores sobre el progreso de los sujetos en el GE fue de 4,78, mientras que tan sólo fue de 2,59 para los sujetos del GC. El contraste mediante la prueba citada anteriormente arroja diferencias muy significativas entre las medias, con un valor *t* de 3,134 y una probabilidad asociada de 0,003 para 52 gl. Por último, las calificaciones escolares también establecieron diferencias muy significativas entre los grupos; mientras que el GE obtenía 5,74 puntos de media en sus calificaciones, el GC

obtenía un promedio de 4,64. El valor t para el contraste efectuado, en el que no se asumió la igualdad de varianzas, fue de 3,291 y la probabilidad asociada de 0,002 para 40,7 grados de libertad. Obviamente, la mayor observación y atención a las notas finales en secundaria se ha derivado de la mayor coherencia observada y analizada anteriormente en esta etapa entre las calificaciones y las valoraciones del profesorado sobre el rendimiento en el área de Lengua. Por tanto, tanto las valoraciones del profesorado sobre el rendimiento escolar, como las notas escolares en el área de Lengua, parecen otorgar un mayor grado de crecimiento significativo —expresado en la valoración sobre el progreso en cada uno de los criterios contemplados— y de rendimiento al grupo al que fue aplicado el dispositivo de evaluación dinámica EDPL.

Análisis de regresión en la prueba ECOS

Al igual que hicimos para las puntuaciones obtenidas en primaria, hemos procedido a realizar diversos análisis de regresión con el fin de complementar la información sobre el valor predictivo del dispositivo de evaluación dinámica en esta etapa. El contenido de estos análisis será similar a los ya presentados para el grupo de primaria y han sido efectuados sobre los mismos elementos: (a) rendimiento escolar en Lengua, medido con la valoración del tutor; (b) rendimiento escolar en Lengua, medido con la valoración del profesorado de apoyo; (c) rendimiento escolar medido con las calificaciones finales en el área de Lengua; y (d) puntuación postest. Hemos procedido igualmente a realizar los análisis de regresión en función de los grupos de sujetos *mejoradores* y *no mejoradores*.

Con relación al rendimiento académico valorado por el profesorado tutor, ya pudimos comprobar que las correlaciones obtenidas con las puntuaciones consideradas no mostraban mucha fuerza. Cuando son introducidas todas las variables en el modelo, sin tener en cuenta el grado de significatividad de los contrastes, la ecuación resultante excluye a la puntuación postest del modelo, cuya predicción conjunta es del 32,8%, resultando significativo el contraste con una probabilidad asociada de $p < 0,043$. En cambio, el resultado del análisis de regresión por pasos muestra que tan sólo la puntuación dinámica EDPL entra en la ecuación, indicando por tanto la significatividad de su contribución al valor que se pretende explicar. El valor R^2 para esta puntuación fue de 0,182, con una F de 4,884 y una significación asociada de 0,038 para 23 grados de libertad. En cuanto a las variables excluidas del modelo, todas correlacionaron positivamente con el rendimiento escolar, aunque en menor grado. La puntuación pretest obtuvo una R^2 de 0,089 y la postest una R^2 de 0,091; ninguno de los contrastes efectuados resultaron significativos, aunque los valores de probabilidad estuvieron cercanos: $p < 0,057$ para el pretest y $p < 0,061$ para el postest. Por su parte, la puntuación de mejora obtuvo una R^2 de 0,009. De esta forma, aún cuando las puntuaciones pretest, postest, y de mejora, obtuvieron correlaciones positivas con el rendimiento escolar, cercanas las dos primeras a la significatividad estadística, fue la puntuación dinámica la única que conseguía explicar y predecir de forma significativa los valores del rendimiento. En síntesis, además de las puntuaciones pretest y postest, la contribución de la puntuación que reflejaba la actuación del sujeto en el dispositivo EDPL resultó esencial para la explicación del rendimiento en educación secundaria.

Por otra parte, el análisis efectuado para predecir la valoración sobre el progreso efectuada por los tutores, mostró igualmente que sólo la puntuación dinámica entraba en la ecuación. El valor de correlación era de $r = 0,510$ y la R^2 de 0,26. El contraste realizado para evaluar el ajuste del modelo fue significativo, con un valor F de 7,745 y una

probabilidad asociada de 0,011. El resto de variables quedó excluido del modelo, siendo sus coeficientes de correlación bajos o muy bajos ($r=0,199$ para el pretest; $r=0,171$ para la puntuación postest; y $r=0,016$ para la mejora). Ninguno de los contrastes efectuados para estas variables resultó significativo.

En cuanto al rendimiento escolar valorado por el profesorado aplicador del dispositivo, el análisis de regresión por pasos efectuado sólo mostró a la puntuación dinámica EDPL como integrante de la ecuación; el valor de correlación entre ambas variables fue de 0,539 y la R^2 de 0,291. El ajuste del modelo resultante fue significativo estadísticamente, con un valor F de 6,142 y una probabilidad asociada de 0,026. Las variables excluidas del modelo (puntuación pretest, postest y mejora) obtuvieron valores de correlación muy bajos ($r=0,100$ para el pretest; $r=0,091$ para el postest y $r=0,047$ para la puntuación de mejora. Cuando se utilizó el procedimiento de introducir todas las variables en bloque, obtuvimos que la puntuación dinámica, la postest y la pretest, explicaban conjuntamente el 31,1% del rendimiento. El incremento en R^2 al introducir el bloque de variables fue apenas de un punto; en cambio, el contraste no resultó significativo ($F = 1,952$; $p < 0,171$). La puntuación de mejora quedó excluida también de esta ecuación.

La valoración del progreso efectuada por el profesorado aplicador fue predicha también de forma significativa por la puntuación dinámica EDPL. El valor de correlación de Pearson fue de 0,698 y la R^2 de 0,487. La puntuación dinámica explicaba por tanto el 48,7% del progreso observado en los sujetos experimentales. La prueba de ajuste del modelo mostró un valor F de 14,252 y una significación de 0,002. El resto de variables correlacionó negativamente con la valoración del progreso; la puntuación pretest, con un valor $r=-0,240$, entró en la ecuación produciendo un incremento en el valor de R^2 de 0,157, con lo que las variables explicaban conjuntamente el 64,4% del progreso observado en los sujetos. El modelo que incluía la puntuación pretest resultó ser también significativo estadísticamente, con un valor F de 12,686 y $p < 0,001$.

Por su parte, las calificaciones escolares también fueron sometidas a un análisis de regresión por pasos para evaluar el valor predictivo de los valores considerados. Este análisis mostró que ninguna de las variables predictoras entraba en la ecuación al no sobrepasar el criterio de significatividad. El valor de correlación más alto lo mostró la puntuación dinámica EDPL ($r=0,306$; sig. 0,076), seguido de la puntuación pretest ($r=0,278$; sig. 0,094). Por otra parte, el postest mostró un valor bajo de correlación y la puntuación de mejora correlacionó negativamente con las notas finales. Cuando se utilizó el procedimiento de introducir todas las variables en bloque, la puntuación postest quedó excluida de la ecuación y los resultados mostraron que el conjunto de variables incluidas explicaba el 22,2% de la varianza, no llegando a ser significativo el contraste ($F = 1,905$; $p < 0,161$).

El análisis de regresión efectuado para predecir la puntuación obtenida en el postest mostró la inclusión de dos variables en la ecuación; en primer lugar, fue introducida la puntuación pretest, con un valor de correlación de $r=0,791$ y una R^2 de 0,626. En segundo lugar la puntuación de mejora, con un valor de correlación de $r=0,598$, ocasionando un incremento en R^2 de 0,374. Entre ambas variables, pretest y mejora, al igual que ocurriera en primaria, explicaban el 100% de los resultados en el postest. El ajuste del modelo resultó muy significativo ($F = 38,461$; $p < 0,000$). Por su parte, la puntuación dinámica correlacionó negativamente con el postest ($r=-0,159$).

Por último, nuestros análisis se dirigieron a tratar de explicar la puntuación de mejora para el grupo de educación secundaria. En este caso, los resultados mostraron que la puntuación postest, con un valor de correlación de $r=0,598$ y una R^2 de $0,357$, explicaba por sí sola de forma muy significativa la diferencia obtenida entre las dos aplicaciones de la prueba criterio. El valor de la prueba F fue de $12,773$ y la probabilidad asociada de $0,002$. Sin embargo, en la ecuación aparecen también las puntuaciones pretest y dinámica; ambas puntuaciones correlacionaron negativamente con la mejora ($r=-0,018$ para el pretest y $r=-0,048$ para la puntuación dinámica), pero su inclusión ocasionó un incremento en el valor de R^2 de $0,643$, por lo que el conjunto de variables incluidas explicaban el 100% de la varianza observada, es decir, de la puntuación de mejora. Probamos a introducir sólo la puntuación postest y la pretest para el análisis, produciéndose idéntico resultado, es decir, ambas variables explicaban la totalidad de la varianza; sin embargo, cuando introdujimos la puntuación postest y la dinámica, ésta última no fue incluida en la ecuación, permaneciendo sólo la postest como factor explicativo. Por tanto, al contrario de lo que ocurriera en primaria, la puntuación dinámica no parece haber tenido parte en la explicación de la puntuación de mejora observada.

El análisis de regresión para los grupos de mejoradores y no mejoradores

Al igual que en educación primaria, hemos procedido a realizar los análisis de regresión en función de los grupos de sujetos experimentales que mejoraron sus puntuaciones tras la aplicación del tratamiento, en este caso, en la prueba de comprensión ECOS; así como en función del grupo que no obtuvo ganancias entre la fase inicial y la final. El grupo de sujetos *mejoradores* estuvo compuesto para estos contrastes por 11 sujetos y obtuvo en el pretest una puntuación de $79,387$; mientras que en el postest su puntuación fue de $88,988$. Su puntuación dinámica EDPL fue de $106,794$; la puntuación de mejora fue de $9,601$; y, por último, la valoración media de los tutores fue de $2,35$ puntos.

El análisis de regresión por pasos para el rendimiento escolar valorado por los tutores mostró que ninguna de las variables era incluida en la ecuación. La puntuación de mejora aparecía como el factor que mejor predecía la valoración académica, obteniendo una correlación de $r=0,516$, con una significación asociada cercana a la significatividad ($p<0,052$). En cuanto a la puntuación dinámica, ésta obtuvo una correlación positiva con el rendimiento ($r=0,397$), al igual que la puntuación postest, con un valor de $r=0,338$. Por su parte, la puntuación pretest correlacionó negativamente ($r=-0,077$) con el rendimiento académico. El análisis individual del valor predictivo de las distintas puntuaciones revelaba que la puntuación de mejora explicaba por sí sola el 26,6% de la varianza observada, produciendo un valor F de $3,269$ y una significación asociada de $0,104$. Con relación a la puntuación dinámica, ésta explicaba el 15,8% del rendimiento, con un cambio en el valor F de $1,686$ y una probabilidad asociada de $p<0,266$. Cuando fueron introducidas en bloque todas las variables, lograban explicar conjuntamente el 42,2% de la varianza, aunque el contraste del modelo resultante no era significativo ($F = 1,705$; $p<0,252$).

El grupo de no mejoradores estuvo compuesto tan sólo por 1 sujeto; su puntuación pretest fue de $79,519$ y la postest de $78,072$; su puntuación dinámica se situó en $96,394$ y la puntuación de mejora fue de $-1,447$. La valoración promedio del tutor fue de $1,429$. Obviamente, no es posible el cálculo de los estadísticos. Hemos procedido a realizar estos cálculos ampliando el subgrupo de no mejoradores, incluyendo, al igual que ya

hicimos para primaria, a todos los sujetos que obtuvieron pérdidas postest, de forma independiente a su puntuación inicial. De esta forma, son 4 sujetos los sujetos que integrarían este subgrupo; la puntuación promedio del pretest fue de 94,348 y la postest de 90,008; su puntuación dinámica se situó en 104,802 y la puntuación de mejora fue de -4,340. La valoración promedio de los tutores fue de 2,571. Una primera valoración de estos promedios —teniendo en cuenta que se trata de una muestra muy reducida— nos hace detenernos en la puntuación media de valoración del rendimiento otorgada a estos sujetos por los tutores, así como en el promedio obtenido en la puntuación dinámica; en efecto, aunque sus puntuaciones en la prueba criterio indicaron pérdidas postest, la valoración del rendimiento ha estado ligeramente por encima de la de sus compañeros que obtuvieron mejoras, e igualmente sus puntuaciones dinámicas no difieren significativamente de las conseguidas por el grupo de mejoradores. Una posible explicación de la inconsistencia de estos resultados apuntaría a que los sujetos de este pequeño grupo también se habrían beneficiado de la aplicación del dispositivo, obteniendo valoraciones positivas de los aplicadores sobre su actuación en las sesiones del mismo, y, así mismo, valoraciones positivas de su rendimiento escolar por parte de sus tutores; sin embargo, su ejecución de la prueba post-tratamiento no habría reflejado este rendimiento, bien por la inadecuación de la prueba estática para mostrarlos, bien por la existencia de un efecto techo en las puntuaciones conseguidas por este grupo, ya que sus puntajes iniciales estaban casi al mismo nivel de los que obtuvieron finalmente los sujetos que obtuvieron ganancias.

El análisis de regresión por pasos no introdujo ninguna variable en la ecuación, a pesar de que los valores de correlación observados para algunas de ellas son muy elevados ($r=0,800$ para la puntuación dinámica EDPL; $r=-0,895$ para la puntuación de mejora; $r=0,016$ para el postest; y $r=0,300$ para la puntuación pretest). Ninguno de estos valores resulta significativo estadísticamente, probablemente debido a la escasa potencia de la muestra considerada. Cuando se optó por introducir las variables en bloque, lograban explicar el 100% de la varianza, aunque no se mostraron valores de contraste del modelo para evaluar su significatividad. En cualquier caso, la introducción parcial de algunas de estas variables no conseguía nunca superar el criterio de significatividad estadística. En síntesis, la puntuación dinámica es la que mejor predice el rendimiento escolar en este pequeño grupo; por otra parte, la puntuación de mejora —caracterizada por pérdidas en este caso—, correlaciona negativamente con el rendimiento debido a que, como hemos podido comprobar, la valoración de los tutores también ha sido positiva para estos sujetos.

En cuanto a la predicción de la valoración del progreso efectuada por los tutores para el grupo de mejoradores, fue también la puntuación dinámica la que mejor predecía el progreso de los sujetos que obtuvieron ganancias postest. El promedio de progreso se situó para este grupo en 4,91 puntos. El valor de correlación fue de $r=0,785$ y la R^2 de 0,616. La prueba F para el contraste del modelo mostró un valor significativo de 14,448 y la probabilidad asociada fue de $p<0,004$. El resto de variables quedó excluido del modelo, con valores de correlación negativos o moderados ($r=-0,156$ para el pretest; $r=0,198$ para el postest; y $r=0,375$ para la puntuación de mejora). En cuanto al valor promedio de progreso obtenido por el sujeto integrante del subgrupo de no mejoradores, éste fue de 3,000. Por su parte, los cálculos para los 4 sujetos que obtuvieron pérdidas, revelan una valoración promedio de progreso de 5,50; por tanto, también superior al observado para el grupo de mejoradores. De igual forma que para la valoración del rendimiento, ninguna de las variables fue introducida en la ecuación del análisis de

regresión por pasos. Los valores de correlación observados fueron: $r=0,546$ para la puntuación dinámica; $r=0,569$ para el pretest; $r=0,331$ para el postest; y $r=-0,680$ para la puntuación de mejora. Al igual que ocurriera con el rendimiento, la valoración del profesorado sobre el progreso del grupo de no mejoradores no estuvo acorde con su actuación en las pruebas criterio. De este modo, volvió a darse una correlación negativa entre la puntuación de mejora y la valoración del progreso. Por otra parte, la puntuación dinámica y la pretest obtuvieron correlaciones bastante altas, que sin embargo y debido probablemente a lo exiguo del grupo, no resultaron significativas. Cuando fueron introducidas todas las variables en la ecuación mediante el método *Introducir*, la explicación lograda llegó al 100% de la varianza, aunque no se mostraron los valores F para evaluar la significatividad del modelo; en cualquier caso, como ya ocurriera en el caso anterior, ninguno de los contrastes parciales efectuados logró que el modelo resultante fuera significativo. Por ejemplo, las puntuaciones dinámicas y pretest conjuntamente explicaban el 91,4% del progreso valorado; sin embargo, la probabilidad asociada al valor de la prueba F fue de $p < 0,293$.

El rendimiento escolar valorado por el profesorado aplicador para el grupo de mejoradores ($N = 10$), también fue sometido al análisis de regresión por pasos. La valoración media obtenida por este grupo fue de 2,39 puntos. La puntuación pretest fue de 79,519, mientras que la postest se situó en 89,501; la puntuación dinámica fue de 108,113 y la puntuación de mejora fue de 9,982. El análisis efectuado no mostró la inclusión de ninguna variable en la ecuación. La correlación más alta, cercana a la significatividad estadística, fue la obtenida por la puntuación de mejora ($r=0,488$; sig. 0,076), seguida por la puntuación dinámica ($r=0,357$; sig. 0,156). Por otra parte, la puntuación postest correlacionó $r=0,300$ con el rendimiento, mientras que el pretest obtuvo una correlación negativa para este grupo ($r=-0,109$). Cuando se procedió a introducir el conjunto de variables en la ecuación, el modelo resultante explicaba el 39,5% del rendimiento valorado por los aplicadores, aunque el contraste para el evaluar el ajuste no resultó significativo ($F = 1,306$; $p < 0,356$).

Tan sólo 2 sujetos del grupo de no mejoradores ampliado recibieron una valoración del rendimiento escolar por parte de los aplicadores. Esta reducida muestra no nos ha permitido extraer conclusiones relevantes sobre las valoraciones efectuadas ni, así mismo, sobre las correlaciones observadas. Tan sólo destacar la continuidad de la tendencia observada en cuanto al promedio de valoración otorgado a los no mejoradores, que en este caso fue de 2,571 puntos, también superior al otorgado para el grupo de mejoradores. También la puntuación dinámica es muy alta, 114,526, lo que unido al dato anterior volvió a situar a la puntuación de mejora en dirección contraria a la valoración del rendimiento efectuada por los aplicadores. Por su parte, la valoración otorgada al sujeto que cumplía las condiciones del subgrupo de no mejoradores (pretest ≤ 84), fue de 1,429.

Con relación a la valoración del progreso efectuada por el profesorado aplicador, valoración que se situó en 5,00 puntos para el grupo de mejoradores, el análisis de regresión por pasos mostró la inclusión de una sola variable en la ecuación; esta fue la puntuación dinámica EDPL, con un valor de correlación $r=0,787$ y una R^2 de 0,619; esta variable predictora lograba explicar de forma muy significativa la varianza observada, con un valor F de 13,024 y una probabilidad asociada de 0,007. En cuanto a la puntuación pretest, al igual que ya ocurriera para la valoración del rendimiento, la correlación fue negativa ($r=-0,177$). Por su parte, la puntuación postest correlacionaba

positivamente con el progreso ($r.0,172$) y la puntuación de mejora obtuvo un valor medio de correlación positiva ($r.0,356$). En este caso, la puntuación de mejora, aunque no de forma significativa, fue en la misma dirección que la valoración sobre el progreso de los sujetos. En este sentido, la lógica apuntaba que el grupo que obtuvo mejoras posttest recibiría mejores valoraciones sobre el progreso en los criterios contemplados.

Para los no mejoradores, la situación fue similar a la descrita para este grupo con relación a la valoración del rendimiento; 1 sujeto integró este subgrupo, ampliado a 2 cuando no se tuvo en cuenta la puntuación inicial. Los valores de correlación mostrados tan sólo nos informan de la dirección de los mismos. La valoración media del progreso fue de 5,00, permaneciendo constantes el resto de valoraciones sobre las demás puntuaciones contempladas y, al igual que en ocasiones anteriores, la puntuación de mejora fue la única que correlacionó negativamente con el criterio.

En cuanto a las calificaciones escolares, el grupo de mejoradores ($N = 11$) obtuvo un promedio de 5,36. El análisis de regresión mostró que ninguna de las variables contempladas sobrepasó el criterio de entrada ($p < 0,05$) y, por tanto, no fueron introducidas en la ecuación. Los valores de correlación fueron los siguientes: el pretest correlacionó negativamente ($r.-0,453$), al igual que la puntuación posttest ($r.-0,071$); la puntuación dinámica obtuvo un valor de correlación positiva, $r.0,128$; y, por último, la puntuación de mejora también correlacionó positivamente con las notas, $r.0,199$. Cuando fueron introducidas todas las variables en bloque, lograban explicar de forma conjunta el 37,8% de la varianza observada, no ajustándose en cambio el modelo a los criterios estadísticos de significatividad ($F = 1,417$; $p < 0,316$). En cuanto al sujeto que cumplía las condiciones de los no mejoradores, éste obtuvo un valor promedio de calificación de 5,00. El grupo de no mejoradores ampliado ($N = 4$) obtuvo una calificación media de 6,5 puntos, sensiblemente superior a la obtenida por los mejoradores. Este dato ahonda en la hipótesis de que tanto las distintas valoraciones efectuadas sobre el rendimiento y el progreso, como la valoración académica final, han revelado una actuación positiva de los sujetos de este grupo, que sin embargo no parece haber sido recogida en su ejecución de las pruebas criterio. El análisis de regresión por pasos efectuado no mostró la inclusión de ninguna de las variables en la ecuación; el valor de correlación más alto fue el de la puntuación dinámica ($r.0,808$; $p < 0,096$), que por sí sola lograba explicar el 65,4% de la varianza observada, aunque el contraste no resultó significativo ($F = 3,773$; $p < 0,192$). Por otra parte, el resto de variables correlacionó negativamente con las notas ($r.-0,146$ para el pretest; $r.-0,353$ para el posttest; y $r.-0,739$ para la puntuación de mejora). Nuevamente, la puntuación de mejora reflejó de forma clara la falta de concordancia entre la actuación negativa de estos sujetos en las pruebas criterio y la valoración positiva que sobre su rendimiento se hace, en esta ocasión, en las calificaciones finales en el área de Lengua.

El análisis de regresión para determinar que valores predecían mejor las puntuaciones posttest ($N = 11$) reflejó la inclusión de dos variables en el modelo; en primer lugar, la puntuación de mejora, con un valor de correlación muy alto, $r.0,895$, y una R^2 de 0,801; en segundo lugar apareció en la ecuación la puntuación pretest, con un valor de correlación también bastante alto, $r.0,735$, ocasionando un incremento en R^2 de 0,199. Ambas variables explicaban de esta forma el 100% de la varianza, es decir, de la puntuación posttest. La puntuación dinámica, por su parte, obtuvo un valor medio de correlación con el posttest ($r.0,380$); de esta forma, la puntuación dinámica lograría explicar por sí sola el 14,4% de la varianza observada. Por lo que respecta al grupo de

no mejoradores ampliado ($N = 4$), el análisis mostró la inclusión de las mismas dos variables que para el grupo de mejoradores y en el mismo orden; la puntuación pretest correlacionó positivamente con el posttest con un valor de $r.0,956$ y una R^2 de $0,914$; por sí sola, esta variable explicaba de forma significativa la varianza ($F = 21,207$; $p < 0,044$). En segundo lugar, entró en la ecuación la puntuación de mejora, que correlacionó también positivamente con el posttest ($r.0,368$). El incremento en R^2 al introducir este segundo factor fue de $0,086$, con lo que ambas variables explicaban conjuntamente el 100% de la puntuación obtenida en el posttest. La puntuación dinámica estableció una correlación negativa con la puntuación final ($r.-0,581$).

Valor de predicción con relación a la escala APSL

El dispositivo de evaluación dinámica puesto en práctica ha pretendido actuar también sobre procesos relacionados con el ajuste personal-social. Del mismo modo, en nuestro diseño de investigación se ha contemplado la evaluación de estos procesos; uno de los procedimientos de evaluación que hemos utilizado ha sido la aplicación de la escala APSL. Con ésta hemos pretendido acceder al conocimiento metacognitivo sobre aspectos relativos a la motivación, las actitudes, las emociones, o el contexto escolar y familiar; todos ellos implicados de forma significativa en los procesos de lectura. Nos disponemos ahora a valorar en qué medida el trabajo desarrollado con la aplicación del dispositivo EDPL se muestra idóneo para otorgar información sobre estos procesos: (a) en comparación con la información que podríamos obtener de la pasación de una prueba estática, como la escala APSL; (b) con relación a la valoración del rendimiento y del progreso observado en la actitud de los sujetos, ya que uno de los ítems de la plantilla de valoración hacía alusión directamente a este aspecto; y (c) con relación al rendimiento y al progreso observados en general en el alumnado participante. Los datos descriptivos, así como los resultados obtenidos de los contrastes efectuados para el grupo experimental, se muestran en las siguientes tablas.

Tabla 80. Valores y puntuaciones obtenidas por el GE en las diversas medidas contempladas

Medidas observadas	Valores medios	Grupo experimental N)
Puntuación dinámica Ajuste P-S	100,289	60
Puntuación dinámica EDPL	100,120	60
Pretest APSL	97,788	57
Posttest APSL	97,564	57
Puntuación de mejora Post – Pret	0,105	54
Valoración tutor (sobre 4)	2,092	59
Suma valoraciones progreso (sobre 7)	3,71	59
Valoración apoyo (sobre 4)	2,192	41
Suma valoraciones progreso (sobre 7)	4,63	41
Calificaciones Lengua (sobre 10)	5,20	46

Tabla 81. Correlaciones entre los valores y puntuaciones obtenidos por el GE en las diversas medidas contempladas.

Los tamaños de las muestras de correlación están en función de los recogidos en la tabla anterior

	Dinám. Ajuste	Dinám. EDPL	Pretest APSL	Postest APSL	Mejora	Valora tutor	∑ valora progreso	Valora apoyo	∑ valora progreso	Notas Lengua
P. dinám. Ajuste P-S		0,818**	0,075	0,132	0,058	0,292*	0,148	0,324*	0,232	0,062
P. dinám. EDPL	0,818**		0,169	0,145	0,030	0,576**	0,466**	0,479**	0,274	0,254
Pretest APSL	0,075	0,169		0,680**	-0,349**	0,187	0,268*	0,229	0,351*	0,108
Postest APSL	0,132	0,145	0,680**		0,450**	0,228	0,192	0,189	0,247	0,333*
P. Mejora Post – Pret	0,058	0,030	-0,349**	0,450**		0,137	-0,012	0,039	-0,030	0,296
Valora. tutor (1-4)	0,292*	0,576**	0,187	0,228	0,137		0,733**	0,693**	0,348*	0,586**
∑ valora. Progreso	0,148	0,466**	0,268*	0,192	-0,012	0,733**		0,520**	0,471**	0,681**
Valora. apoyo (1-4)	0,324*	0,479**	0,229	0,189	0,039	0,693**	0,520**		0,627**	0,436**
∑ valora. Progreso	0,232	0,274	0,351*	0,247	-0,030	0,348*	0,471**	0,627**		0,364*
Notas Lengua	0,062	0,254	0,108	0,333*	0,296	0,586**	0,681**	0,436**	0,364*	

La inspección de los datos provenientes de los análisis correlacionales nos muestra en primer lugar una altísima correlación entre la puntuación dinámica total obtenida en el dispositivo EDPL y la puntuación dinámica obtenida en los procesos de ajuste personal-social. Unido a este dato, que nos habla de la coherencia entre la valoración otorgada por los aplicadores con relación al dispositivo en su conjunto y la valoración realizada sobre los procesos de ajuste que forman parte del mismo, al tiempo que informa sobre la relación estrecha que hemos encontrado entre el rendimiento de los sujetos en el dispositivo y la puesta en práctica de habilidades relativas a estos procesos de ajuste social y personal; nos encontramos que la puntuación dinámica para los procesos de ajuste también correlaciona positiva y significativamente con la valoración de los tutores sobre el rendimiento y más aún con la valoración sobre el rendimiento que efectúan los aplicadores. Por el contrario, esta puntuación dinámica parcial obtiene una muy baja correlación con el pretest de la escala APSL, así como con la puntuación de mejora de esta prueba. Con el postest el nivel de correlación sube, aunque no llega a ser significativo. Estos datos, unidos a los que expresan un menor grado de correlación entre las puntuaciones estáticas pretest y postest, y las puntuaciones de rendimiento escolar, nos indican que la puntuación dinámica obtenida de la valoración efectuada sobre la puesta en práctica de procesos de ajuste personal-social durante la realización del dispositivo, predice el rendimiento escolar significativamente y mejor que las puntuaciones obtenidas en una escala de evaluación estática como APSL.

Por otra parte, ya analizamos comparativamente con las pruebas estáticas las correlaciones que mantenía la puntuación dinámica global con las valoraciones de los tutores o del profesorado de apoyo. Esos datos estaban limitados a las muestras de primaria o de secundaria, en función de la prueba de comprensión analizada. En este caso, la correlación mostrada entre la puntuación EDPL y las valoraciones del profesorado engloban al alumnado de primaria y de secundaria que realizó la prueba APSL. Como podemos observar, las correlaciones son muy significativas con la

valoración de los tutores sobre el rendimiento escolar y sobre el progreso del alumnado, así como con la valoración del rendimiento efectuada por el profesorado de apoyo. Son también positivas y de valor medio, aunque sin llegar a ser significativas, para el progreso valorado por los profesores de apoyo y para las calificaciones. Por su parte, las puntuaciones pretest de la escala APSL obtienen correlaciones significativas ($p < 0,05$) con la valoración de progreso efectuada por los tutores y por el profesorado de apoyo, mientras que para el resto de valoraciones obtienen un nivel de correlación medio o bajo. El postest obtiene correlaciones medias o bajas, que no llegan a la significatividad estadística salvo con las calificaciones, al nivel 0,05. La puntuación de mejora obtiene índices muy bajos de correlación, salvo para las calificaciones finales igualmente. En síntesis, podríamos decir que la puntuación dinámica EDPL muestra un nivel predictivo sobre el rendimiento escolar muy superior al otorgado por las puntuaciones que se derivan de las aplicaciones pretest y postest de la escala APSL, así como del ofrecido por la puntuación de mejora en esta prueba.

Hemos calculado en qué medida la puntuación dinámica global y la obtenida de los procesos de ajuste personal-social, predicen la valoración efectuada por el profesorado sobre el ítem de la plantilla que hacía referencia explícita a la actitud ante la lectura. La puntuación media obtenida para la actitud en función de la valoración efectuada por los tutores y por el profesorado de apoyo ha sido idéntica, con un valor de 2,54 puntos. En cuanto al progreso valorado por los tutores, el promedio se ha situado en 0,68 (sobre 1 ítem) y el valorado por el profesorado de apoyo ha sido de 0,73. Los datos del contraste los mostramos en la siguiente tabla.

Tabla 82. Correlaciones entre las valoraciones efectuadas sobre la actitud, y los valores y puntuaciones obtenidos por el GE en las diversas medidas contempladas.

Los tamaños de las muestras de correlación están en función de los recogidos en la tabla anterior

	Dinám. Ajuste	Dinám. EDPL	Pretest APSL	Postest APSL	Mejora	Actitud tutor	∑ valora Progreso	Actitud apoyo	∑ valora progreso	Notas Lengua
Actitud tutor (1-4)	0,397**	0,504**	0,106	0,164	0,090		0,443**	0,561**	0,139	0,345*
∑ valora. Progreso	0,187	0,276*	0,224	0,032	-0,228	0,443**		0,298	0,352*	0,446**
Actitud apoyo (1-4)	0,474**	0,457**	0,314*	0,295	0,005	0,561**	0,298		0,495**	0,391*
∑ valora. Progreso	0,245	0,119	0,325*	0,169	-0,166	0,139	0,352*	0,495**		0,280

Al igual que para el rendimiento escolar en su conjunto, para la actitud ante la lectura, tomada de forma individual, las puntuaciones dinámicas, tanto la obtenida de la valoración de los procesos de ajuste personal-social, como la global, muestran niveles de correlación muy significativos. Esto ocurre con relación a la valoración que tutores y profesorado de apoyo efectúan sobre la actitud; también correlaciona positiva y significativamente al nivel 0,05 la puntuación dinámica global con la valoración de progreso efectuada por los tutores. Por su parte, la puntuación pretest en APSL correlaciona positivamente con las valoraciones sobre rendimiento y progreso de los tutores, llegando a ser estadísticamente significativas al nivel 0,05 con las valoraciones sobre rendimiento y progreso que efectúan los aplicadores. La puntuación postest APSL no obtiene niveles más altos de correlación que el pretest salvo para la valoración sobre el rendimiento realizada por los tutores, y la puntuación de mejora obtiene correlaciones muy bajas para las valoraciones sobre rendimiento y negativas para las valoraciones

sobre el progreso del alumnado. Por último, es destacable la relación existente entre las valoraciones efectuadas por el profesorado tutor y el de apoyo con relación a la actitud ante la lectura; la correlación es muy significativa para la valoración sobre el rendimiento y significativa al nivel 0,05 para el progreso. También se muestran significativas o muy significativas las correlaciones con las calificaciones escolares. En definitiva, podemos concluir que la puntuación dinámica global y la puntuación dinámica relativa a los procesos de ajuste personal-social se han mostrado más relacionadas con la actitud del alumnado ante la lectura que la puntuación obtenida en las aplicaciones pretest y posttest de la escala APSL.

Hemos realizado los cálculos de las correlaciones para el GC tomado en su totalidad. La valoración promedio del tutor sobre el rendimiento escolar ($N = 54$) fue de 1,876 y el valor medio de la suma sobre las valoraciones positivas de progreso fue de 2,46; así mismo, la valoración media del profesorado de apoyo ($N = 7$) fue de 1,735 y la valoración media del progreso fue de 4,00. El promedio de las calificaciones escolares ($N = 44$) arrojó un resultado de 4,43. Por otra parte, la puntuación media pretest en la escala APSL ($N = 60$) fue de 95,068, y la posttest ($N = 56$) fue de 93,413. La puntuación de mejora ($N = 56$) se situó en -2,223 puntos. Los valores muestran una diferencia inicial en la valoración media que recibieron los sujetos del GE y la que obtuvieron los sujetos del GC. Todas las valoraciones se sitúan por debajo para este último grupo, incluidas las calificaciones escolares en el área de Lengua.

Los contrastes efectuados revelan que, al igual que ya ocurriera para el GE, la puntuación pretest y la posttest de la escala APSL correlacionan muy significativamente ($r.0,557$). También correlacionan muy significativamente la puntuación de mejora y el posttest de la escala ($r.0,481$), mientras que el pretest y la puntuación de mejora mantienen una correlación negativa y también muy significativa ($r.-0,461$). Por otra parte, el pretest correlaciona de forma muy significativa con la valoración del tutor sobre el rendimiento ($r.0,400$) y sobre el progreso del alumnado ($r.0,511$), mientras que el posttest obtiene niveles medios de correlación, que no llegan a ser significativos para la valoración del tutor sobre el rendimiento, y la puntuación de mejora obtiene valores negativos de correlación. Algunos de estos últimos datos son los que permiten establecer una comparación relevante con los anteriormente expuestos para el GE. En primer lugar, la mayor significatividad de la correlación de la puntuación pretest con el rendimiento escolar valorado por los tutores, con relación a la obtenida por la puntuación posttest, hace pensar que en el grupo control la valoración inicial se ha mostrado suficiente para informar sobre el rendimiento de los sujetos; este dato se vería también corroborado por el mayor valor de la correlación del pretest con las calificaciones escolares ($r.0,159$ para el pretest y $r.0,088$ para el posttest). De esta forma, la ausencia de tratamiento habría hecho válida la predicción inicial del pretest sobre el rendimiento. Sin embargo, para el GE, la aplicación del dispositivo EDPL habría introducido un elemento nuevo y significativo que habría desbancado a la puntuación pretest como la variable que mejor predice el rendimiento, situándose en su lugar. En efecto, como hemos podido ver, tanto la puntuación dinámica global ($r.0,576$) como la relativa a los procesos de ajuste ($r.0,292$) correlacionaban con el rendimiento escolar valorado por los tutores significativamente por encima de la puntuación pretest ($r.0,187$). Esto también resultaría válido para la valoración del rendimiento efectuada por el profesorado aplicador.

Análisis de regresión en la escala APSL

El análisis de regresión efectuado para el GC con relación al rendimiento escolar, nos muestra que la puntuación pretest figura como la única variable que es introducida en la ecuación, con un valor R^2 de 0,213; el contraste resulta muy significativo, con una $F = 11,128$ y una $p < 0,002$. Por su parte, para el GE, la ecuación contempla sólo las dos puntuaciones dinámicas, excluyendo del modelo a las restantes variables. En primer lugar fue introducida la puntuación dinámica global, con un valor R^2 de 0,309, un valor F de 22,763 y una probabilidad asociada de 0,000; en segundo lugar la puntuación dinámica de los procesos de ajuste, que ocasionó un incremento en el valor de R^2 de 0,082. Entre ambas variables explicaban el 39,1% del rendimiento escolar valorado por los tutores, resultando muy significativo el contraste del modelo ($F = 16,048$; $p < 0,000$).

Con relación al progreso valorado por los tutores obtenemos idéntico resultado. El análisis realizado para el GC muestra una ecuación con la puntuación pretest como única variable predictora. El valor de R^2 fue de 0,369, la F resultante de 23,986 y la probabilidad asociada de 0,000. Por su parte, el análisis para el GE mostró nuevamente una ecuación compuesta por las dos puntuaciones dinámicas. En primer lugar la puntuación dinámica global, con un valor R^2 de 0,191, valor que incrementaba la introducción de la puntuación dinámica de los procesos de ajuste en 0,148. Entre las dos variables explicaban el 33,9% del progreso valorado, resultando de nuevo muy significativo el modelo ($F = 12,811$; $p < 0,000$).

La valoración del rendimiento por parte del profesorado aplicador en el GE ($N = 39$) también fue objeto de un análisis de regresión por pasos con objeto de evaluar la potencia de la predicción de las distintas variables consideradas. Los resultados mostraron que sólo la puntuación dinámica EDPL era introducida en la ecuación, con un valor R^2 de 0,224; el modelo compuesto por este único factor era significativo, con un valor F de 10,653 y una probabilidad asociada de 0,002. El resto de variables fue excluido del modelo al no cumplir los criterios de entrada. Por su parte, el análisis para el GC ($N = 7$) mostró que la puntuación posttest era la que mejor predecía el rendimiento valorado por el profesorado de apoyo, con un valor para R^2 de 0,745 ($F = 14,614$; $p < 0,012$).

El análisis de regresión por pasos, efectuado para la valoración del progreso realizada por el profesorado aplicador en el GE ($N = 39$), mostró en este caso que era la puntuación pretest la única que era introducida en la ecuación. El valor R^2 no fue muy alto: 0,102, obteniendo una F para el contraste del modelo de 4,220 con una probabilidad asociada de $p < 0,047$. Cuando se introdujeron en bloque todas las variables contempladas, el conjunto lograba explicar el 16,7% del progreso valorado, aunque el modelo resultante no cumplía los parámetros de significatividad estadística ($F = 1,700$; $p < 0,173$). Por otra parte, la introducción en bloque de las dos puntuaciones dinámicas (global y relativa a los procesos de ajuste), explicaba el 7,5% de la valoración efectuada sobre el progreso, aunque obviamente tampoco resultó significativo el contraste del modelo. Para el GC ($N = 7$) el análisis efectuado no mostró ninguna variable que lograra salvar los criterios de entrada en la ecuación de regresión. Cuando fueron introducidas en bloque todas las variables consideradas, lograban explicar el 30,8% del progreso valorado, aunque el modelo no resultaba significativo ($F = 2,334$; $p < 0,213$), probablemente debido a lo exiguo de la muestra.

Con relación a las calificaciones finales en el área de Lengua, el análisis de regresión por pasos para el GE (N = 44) reveló inicialmente a la puntuación obtenida en el postest como la única integrante de la ecuación, explicando el 11,1% del rendimiento. El contraste resultaba significativo con un valor F de 5,260 y una probabilidad asociada de 0,027. Sin embargo, introduciendo conjuntamente las variables dinámicas (puntuación dinámica global + puntuación dinámica relativa a los procesos de ajuste), el modelo resultante explicaba el 14,6% del rendimiento valorado mediante las calificaciones, resultando significativo estadísticamente (F = 3,669; $p < 0,034$). También resultó significativo el modelo compuesto por la puntuación dinámica global y la puntuación de mejora; en este caso, la R^2 fue de 0,141, obteniendo el valor F para el contraste 3,353 y la significación asociada a dicho valor 0,045. Por su parte, el análisis efectuado para el GC (N = 34) no mostró factores significativos que explicaran el rendimiento. Al introducir conjuntamente las variables consideradas (pretest, postest y mejora), éstas lograban explicar el 2,2% de la varianza, con un valor F de 0,350 y una probabilidad de obtener un resultado menor o igual a éste de 0,708.

Hemos realizado el mismo análisis para la puntuación postest; éste muestra para el GE (N = 54) que el pretest y la puntuación de mejora explican en su totalidad la puntuación postest. La ecuación resultante introduce en primer lugar a la puntuación pretest con un valor R^2 de 0,462, resultando significativo el contraste con una F de 44,655 y una $p < 0,000$. En segundo lugar es introducida la puntuación de mejora, incrementando el valor R^2 en 0,538 hasta llegar a explicar conjuntamente ambas variables el 100% de la puntuación final. Tanto la puntuación pretest como la puntuación de mejora explicarían por sí solas de forma significativa la varianza observada. Por su parte, las puntuaciones dinámicas explicaban sólo el 2,2% del resultado final, pero conjuntamente con la puntuación de mejora lograban explicar el 22,7% del postest, resultando significativo el modelo (F = 4,904; $p < 0,005$). Para el GC (N = 56) el análisis efectuado produce similares resultados; la puntuación pretest y la puntuación de mejora, por este orden, son introducidas en la ecuación de regresión por pasos, explicando inicialmente el pretest el 31% del postest e incrementándose el valor R^2 en 0,69 con la introducción de la puntuación de mejora. En este caso, la puntuación de mejora también explicaba de forma significativa la puntuación postest por sí sola.

Por último, en cuanto a la regresión efectuada para el GE con relación a la valoración de la actitud por parte del profesorado, destacar que la puntuación dinámica global fue el factor que mejor predecía la valoración sobre la actitud ante la lectura realizada por los tutores, así como el progreso observado por éstos mismos con relación a la actitud. En ambos contrastes, la puntuación dinámica global apareció como la única capaz de salvar el criterio de entrada en la ecuación de regresión. El valor R^2 para el primero de los análisis mencionados fue de 0,267, resultando significativo el modelo con una F de 18,592 y una probabilidad asociada de 0,000. Por su parte, para la valoración del progreso, el análisis mostró un valor R^2 sensiblemente menor, 0,077, con una F de 4,275 y una probabilidad de 0,044. En cuanto a la valoración realizada por el profesorado aplicador sobre la actitud, el análisis mostró en este caso a la puntuación dinámica relativa a los procesos de ajuste como la única variable introducida en la ecuación, explicando por sí sola y de forma significativa el 23,1% de la actitud (F = 11,099; $p < 0,002$). Por otra parte, el análisis efectuado sobre la valoración realizada por estos mismos profesores sobre el progreso en la actitud ante la lectura, observó un comportamiento similar al ya reflejado para la valoración global sobre el progreso. Aunque inicialmente no fue introducida ninguna variable en la ecuación, fue la puntuación pretest la que mejor explicaba el progreso observado; introducida de forma

individual, obtenía un valor R^2 de 0,106, con una F de 4,491 y una probabilidad asociada de 0,041. Por su parte, las puntuaciones dinámicas introducidas conjuntamente lograban explicar un porcentaje similar del progreso ($R^2 = 0,098$), aunque el contraste del modelo no resultaba estadísticamente significativo. Tampoco resultó significativo el contraste del modelo que introducía en bloque a la totalidad de las variables consideradas, incluyendo la puntuación pretest.

Valor de predicción con relación a la prueba de CI

Aunque en un principio nuestro estudio no se dirigía específicamente al establecimiento de una mejora en el CI de los sujetos experimentales, como fruto de su participación en las sesiones de aplicación del dispositivo de evaluación dinámica, ya comentamos y analizamos en páginas anteriores, que tanto el contenido del mismo como los patrones metodológicos puestos en práctica en su aplicación, habrían contribuido de forma significativa a la obtención de ganancias en la prueba de inteligencia utilizada, especialmente, como consecuencia de la transferencia que se habría producido de las habilidades metacognitivas trabajadas. En este apartado nos disponemos a analizar el poder explicativo que hayan podido tener las distintas puntuaciones dinámicas consideradas, con relación a las puntuaciones posttest obtenidas en la prueba de inteligencia. Este valor predictivo del dispositivo EDPL será analizado en contraste con el poder explicativo de la puntuación pretest de la prueba de CI. En contraposición a una lógica de la inteligencia como constructo que revela una concepción estática de la misma, la lógica dinámica en la que se sustenta nuestro trabajo, contempla la posibilidad de modificar sustantivamente la influencia que la puntuación inicial ejerce sobre la puntuación posttest como fruto del tratamiento aplicado. De esta forma, se esperaba que tanto la puntuación dinámica como la puntuación de mejora contribuyeran adicionalmente a la explicación de la puntuación obtenida por los sujetos en la prueba postratamiento. Mostramos en primer lugar los valores y las correlaciones obtenidas entre las diversas pruebas y puntuaciones; hemos incluido aquí las correlaciones entre las pruebas de comprensión, ajuste e inteligencia. En los análisis subsiguientes podremos mostrar en qué forma la inteligencia ha influido en las puntuaciones obtenidas en cada una de las demás pruebas de criterio utilizadas.

Tabla 83. Valores y puntuaciones obtenidas por el GE en las diversas medidas contempladas

Medidas observadas	Valores medios	Grupo experimental N)
CI pretest	82,57	60
CI posttest	88,00	57
Mejora CI	4,789	57
Puntuación dinámica Metacognitiva	100,215	60
Puntuación dinámica global EDPL	100,120	60
Valoración rendimiento tutor	2,092	59
Valoración progreso tutor	3,71	59
Valoración rendimiento apoyo	2,192	41
Valoración progreso apoyo	4,63	41
Calificaciones Lengua	5,20	46

Tabla 84. Correlaciones entre las puntuaciones CI y los valores y puntuaciones obtenidos por el GE en las diversas medidas contempladas.

Los tamaños de las muestras de correlación están en función de los recogidos en la tabla anterior

Medidas obtenidas	CI pretest	CI postest	Mejora CI	P. dinámica Meta.	P. dinámica global
CI pretest		0,720**	-0,275*	0,032	0,092
CI postest	0,720**		0,469**	0,154	0,168
Mejora CI	-0,275	0,469**		0,172	0,129
Pretest ECO 1	0,384*	0,240	-0,143	-0,500**	-0,503**
Postest ECO 1	0,289	0,311	0,062	-0,180	-0,198
Mejora ECO 1	-0,043	0,040	0,105	0,206	0,228
Pretest ECOS	0,138	0,030	-0,083	-0,349	-0,220
Postest ECOS	0,384	0,349	0,029	-0,301	-0,159
Mejora ECOS	0,274	0,336	0,134	-0,095	-0,048
Pretest APSL	0,200	0,181	0,067	0,156	0,169
Postest APSL	0,196	0,082	-0,117	0,108	0,145
Mejora APSL	0,030	-0,113	-0,203	0,009	0,030
Puntuación dinámica Metacognitiva	0,032	0,154	0,172		0,972**
Puntuación dinámica global EDPL	0,092	0,168	0,129	0,972**	
Valoración rendimiento tutor	0,082	0,147	0,127	0,544**	0,576**
Valoración progreso tutor	0,253	0,306*	0,120	0,434**	0,466**
Valoración rendimiento apoyo	0,219	0,367*	0,285	0,476**	0,479**
Valoración progreso apoyo	0,033	0,239	0,397*	0,301	0,274
Calificaciones Lengua	0,149	0,118	0,026	0,195	0,254

En primer lugar debemos destacar que la puntuación obtenida en la prueba de CI, en la fase pretest, mantiene una alta y muy significativa correlación con la puntuación postest; este resultado responde a lo que cabe esperar de la doble pasación de una prueba de inteligencia general; más aún en un intervalo de tiempo como el que ha transcurrido en nuestro estudio. En segundo lugar, podemos observar que la puntuación de mejora también correlaciona de forma positiva y muy significativamente con el postest, incrementando notablemente su valor de correlación con respecto al obtenido con la puntuación pretest. Estos datos, unidos al de las ganancias observadas en el GE, apuntan a la relevancia que podría haber tenido el tratamiento aplicado entre la realización de las pruebas de criterio. Sin embargo, para confirmar esto, debemos cerciorarnos de que para el GC la puntuación de mejora no obtuvo resultados similares, o lo que es lo mismo, la puntuación pretest habría tenido un peso significativamente mayor en la puntuación obtenida en el postest (estos análisis serán revisados a continuación). Por otra parte, es relevante destacar que tanto la puntuación dinámica

metacognitiva como la global incrementan sus valores de correlación en el postest con respecto a los obtenidos en el pretest. Igualmente, los valores de correlación de estas puntuaciones con la mejora CI también son superiores a los obtenidos por el pretest con la mejora.

En otro sentido, uno de los aspectos fundamentales en nuestro análisis de los resultados relativos al valor de predicción, consiste en comprobar si el rendimiento escolar puede predecirse mejor desde puntuaciones obtenidas con pruebas estáticas o desde las puntuaciones dinámicas. En este sentido, los resultados muestran con nitidez que los niveles de correlación establecidos entre el rendimiento y las dos puntuaciones dinámicas contempladas, superan a los observados para las puntuaciones iniciales de CI, así como también (aunque en menor grado) con relación a las puntuaciones finales. Los valores de correlación obtenidos por las puntuaciones dinámicas logran explicar muy significativamente el rendimiento escolar valorado por los tutores y los profesores de apoyo, así como el progreso de los sujetos valorado por los tutores. De esta forma, podemos concluir que han sido las puntuaciones obtenidas de la valoración del *proceso de aprendizaje* las que mejor se han relacionado con la valoración del rendimiento escolar realizada por el profesorado; así mismo, con relación a las calificaciones escolares en el área de Lengua, también se han obtenido valores de correlación más altos para las puntuaciones dinámicas que para las puntuaciones CI, aunque en este caso, los valores no llegan a ser significativos. En definitiva, la puntuación dinámica ha mostrado una mayor relación con el rendimiento escolar que las puntuaciones estáticas, tanto las obtenidas de pruebas relacionadas con el contenido a evaluar (las pruebas de comprensión y las de ajuste personal-social con relación a la lectura), como las obtenidas de pruebas de inteligencia general.

Con respecto a la relación de las puntuaciones dinámicas con las pruebas criterio de contenido lingüístico, podemos observar que la puntuación de mejora en la prueba de primaria, es la que obtiene valores de correlación más altos con las puntuaciones dinámicas. En efecto, mientras que el pretest y el postest de dicha prueba obtienen correlaciones negativas con los puntajes dinámicos, la puntuación de mejora correlaciona positivamente con un valor medio ($r=0,206$ para la puntuación metacognitiva y $r=0,228$ para la puntuación dinámica global). Una tendencia similar en los resultados se obtuvo con la prueba ECOS; en este caso, los valores de correlación de las puntuaciones dinámicas con la mejora son negativos, al igual que los que mantienen con las puntuaciones iniciales y finales; aunque se observó en este caso una disminución sensible en el valor negativo de correlación, con respecto al obtenido con la puntuación pretest. Por otra parte, con relación a la escala APSL, los valores de correlación con las puntuaciones dinámicas consideradas son positivos, aunque bajos; no observándose niveles de correlación más óptimos para la puntuación de mejora con respecto a la pretest.

También podemos destacar que las puntuaciones CI no han obtenido niveles altos de correlación con el resto de pruebas criterio, y especialmente con las pruebas de comprensión. El valor de la correlación es moderado para la puntuación pretest de ECO 1 y la postest de ECOS, ambas con un valor $r=0,384$; en cambio, para el resto, los valores son medios o bajos. Las medidas pretest y de mejora obtenidas en las pruebas de comprensión obtienen valores de correlación significativamente más altos con la puntuación postratamiento que las medidas de inteligencia general, ya sean éstas iniciales o finales. Posteriormente, en el capítulo de Discusión, abordaremos con mayor

profundidad el análisis de estos resultados. Por último, debemos apuntar la altísima correlación entre la puntuación dinámica global y la relativa a los procesos metacognitivos. Este dato nos habla del peso que estos procesos han tenido en la puntuación dinámica EDPL, así como de la conexión que se habría producido entre la valoración global del proceso de comprensión y la valoración más específica de los procesos metacognitivos implicados en la misma.

Análisis de regresión en la prueba de CI

Los valores correlacionales examinados han mostrado algunos indicios relevantes a favor de algunas de nuestras hipótesis experimentales; sin embargo, aún debemos someter los datos a un análisis de regresión que nos permita determinar con una mayor fiabilidad y precisión en qué medida las puntuaciones observadas en la prueba criterio de inteligencia general pueden predecirse a raíz de las puntuaciones de mejora obtenidas en la misma prueba o de las puntuaciones obtenidas de la valoración del proceso de aplicación dinámica del dispositivo. Por otra parte, también nos permitirán evaluar diferencialmente el valor predictivo de las puntuaciones dinámicas sobre el rendimiento escolar con respecto a las puntuaciones estáticas obtenidas de la prueba de CI.

En primer lugar hemos efectuado un análisis de regresión por pasos para evaluar el peso de las puntuaciones pretest y de mejora en la explicación del postest CI en los grupos experimental y control. Los resultados para el GE ($N = 57$) muestran una ecuación con dos factores, pretest y mejora. La puntuación pretest es introducida en primer lugar, con un valor R^2 de 0,519. El contraste efectuado para evaluar la significatividad de esta variable es significativo, con un valor F de 59,267 y una probabilidad asociada de 0,000. La introducción de la puntuación de mejora ocasiona un incremento en el valor R^2 de 0,481, con lo que el modelo resultante compuesto por ambos factores explica el 100% de la varianza. La puntuación de mejora por sí sola también explica de forma significativa la puntuación postest, obteniendo un valor R^2 de 0,220 ($F = 15,497$; $p < 0,000$). Las puntuaciones dinámicas fueron incluidas en el análisis, sin embargo la escasa fuerza de las correlaciones con el postest no permitieron su introducción en el modelo resultante. La conjunción de la puntuación de mejora y de las puntuaciones dinámicas (metacognitiva y global) explicaría el 24,8% del postest, resultando igualmente significativo el contraste del modelo. Por su parte, las puntuaciones dinámicas introducidas conjuntamente explicarían por sí solas el 3% de la varianza.

El análisis efectuado para valorar en qué medida los resultados anteriores se han producido igualmente en el GC ($N = 42$), nos han revelado que el pretest ha ejercido un peso sensiblemente superior en la puntuación postest. La ecuación de regresión muestra las dos variables, pretest y mejora, aunque el valor R^2 del pretest fue de 0,691 y el de la puntuación de mejora de 0,309. Al igual que para el GE, el contraste efectuado era muy significativo, con una probabilidad asociada de 0,000; sin embargo, el mayor peso del pretest en la explicación del postest, así como el menor incremento en el valor de R^2 como consecuencia de la introducción de la puntuación de mejora, indicarían un mayor peso en el GE de la puntuación que refleja el cambio producido. En efecto, mientras que para el GC el pretest explicaría el 69,1% del resultado en el postest de la prueba de CI, para el GE, el pretest tan sólo explicaría el 51,9%. Esta diferencia en los porcentajes de 17,2 puntos es idéntica obviamente a la observada para la puntuación de mejora: el 30,9% del postest explica la mejora en el GC, mientras que en el GE este porcentaje se acerca al valor de explicación del pretest, 48,1%.

Por otra parte, hemos realizado también este análisis de regresión por pasos para valorar el valor predictivo de las puntuaciones dinámicas sobre el rendimiento escolar con relación al de las puntuaciones CI. La ecuación resultante para el rendimiento valorado por los tutores nos muestra que la puntuación dinámica global es la única que salva el criterio de entrada, con un valor R^2 de 0,324. La F fue de 25,930 y la probabilidad asociada de 0,000. Idéntico resultado obtenemos del análisis efectuado para la valoración del progreso de los sujetos efectuada por los tutores. La puntuación dinámica aparece como la única variable en la ecuación, explicando de forma significativa el 24,1% de la varianza. Para la valoración del rendimiento escolar realizada por el profesorado aplicador, también fue la puntuación dinámica global la que mejor predecía el resultado, con un 24,6% de la varianza explicada. Los tres contrastes derivados de los modelos mencionados, y compuestos tan sólo por un factor, resultaron ser muy significativos estadísticamente y en todos ellos, aunque no era introducida en la ecuación inicialmente, la puntuación dinámica metacognitiva obtenía niveles de correlación similares a los de la puntuación dinámica global. De hecho, cuando esa misma puntuación era introducida en las mismas condiciones pero de forma individual, el resultado del análisis mostraba un porcentaje similar de varianza explicada, resultando así mismo significativo el contraste de evaluación de la significatividad.

Por su parte, el análisis efectuado para la valoración del progreso efectuada por el profesorado aplicador, mostró a la puntuación de mejora CI como la mejor predictora del resultado. El valor R^2 fue en este caso de 0,158 y el contraste resultaba significativo con una F de 6,732 y una probabilidad asociada de 0,014. Para las calificaciones escolares, ninguna variable salvó el criterio de entrada estipulado en el análisis. Cuando fueron introducidas en bloque todas las variables, éstas explicaban de forma conjunta el 8,3% de la varianza; sin embargo, cuando fueron introducidas sólo las puntuaciones dinámicas, el porcentaje de varianza explicada se incrementó hasta el 11%, a pesar de lo cual el modelo resultante no llegaba a ser significativo estadísticamente ($F = 2,655$; $p < 0,082$).

Los análisis de regresión por pasos para el GC mostraron, en ausencia de las puntuaciones dinámicas, que los puntajes obtenidos en el pretest y el postest de la prueba de CI, así como las puntuaciones de mejora, no predijeron de forma significativa el rendimiento escolar. Los resultados revelaron que para ninguno de los contrastes efectuados en función de las distintas valoraciones realizadas por los tutores y los profesores de apoyo sobre el rendimiento y sobre el progreso de los sujetos, así como también para la valoración del rendimiento escolar que se derivaba de las calificaciones escolares en el área de Lengua, fue incluida inicialmente ninguna de las variables contempladas en la ecuación de regresión. En general, las correlaciones observadas del rendimiento escolar con las medidas de CI fueron medias o bajas, llegando a ser negativas en algunos casos.

Por último, cuando hemos procedido al análisis diferencial del valor predictivo de las puntuaciones dinámicas sobre el rendimiento escolar, en función del nivel intelectual de los sujetos ($CI \leq 84$ / $CI > 84$), los resultados nos mostraron, en general, valores de correlación significativos o muy significativos para las distintas valoraciones efectuadas por los tutores y por el profesorado de apoyo, tanto sobre el rendimiento como sobre el progreso observado en los sujetos. Así mismo, estos valores han sido observados con relación a la fase pretest y también a la fase postest de la investigación. Por su parte, los análisis de regresión realizados mostraron que las puntuaciones dinámicas eran

incluidas en la ecuación, prediciendo de forma significativa el rendimiento escolar en la mayoría de los casos. No obstante, se observa una ligera tendencia a la obtención de valores de correlación más bajos en el grupo de nivel intelectual más alto. Este dato queda confirmado, aunque con un número de sujetos exiguo ($N = 6$), cuando se han efectuado estos contrastes con los sujetos que obtuvieron puntuaciones CI que igualaban o superaban los 100 puntos; en estos casos, aunque las puntuaciones dinámicas seguían manteniendo niveles de correlación positivos con el rendimiento, no lograban explicar significativamente el modelo.

GRADO DE EFICACIA

Aunque posteriormente volveremos sobre este aspecto, queremos destacar aquí algunos elementos que consideramos de interés. En el esquema del análisis propuesto, hacíamos referencia al grado de eficacia con el que el procedimiento dinámico se presenta, como uno de los elementos del perfil de cuatro puntos que hemos adoptado en nuestro análisis del dispositivo EDPL. En este caso, el análisis debía dirigirse fundamentalmente a determinar en qué medida aspectos como el tiempo empleado o el esfuerzo invertido en el procedimiento, quedaban justificados en función de la información adicional que podría aportar el proceso de evaluación dinámica presentado con relación a la aplicación de pruebas convencionales. En este sentido, a lo largo del presente capítulo, y especialmente en nuestra valoración de: (a) la calidad de los datos obtenidos y de la fiabilidad de las valoraciones obtenidas en el dispositivo; (b) del tipo de información ofrecida y de las aportaciones metodológicas suplementarias; así como (c) del valor de predicción de las puntuaciones dinámicas sobre el rendimiento académico; hemos hecho alusión a que la información que puede obtenerse de la aplicación del dispositivo EDPL es cualitativamente distinta a la que obtendríamos a través de una prueba convencional de evaluación de la comprensión. Esta información, que podría considerarse «adicional» a la otorgada por las pruebas estáticas, versaría sobre las dificultades de aprendizaje que manifiesta el sujeto durante el proceso de resolución de las tareas, así como sobre aquellos patrones metodológicos de evaluación dinámica que se hubieran revelado determinantes o eficaces en el proceso de aplicación.

Por otra parte, los análisis del valor de predicción del dispositivo dinámico sobre el rendimiento escolar, muestran valores que expresan una contribución muy significativa de las puntuaciones dinámicas, extraídas de la valoración de los profesores-aplicadores sobre la puesta en práctica del dispositivo, como factor explicativo del rendimiento académico. Este poder de predicción se ha mostrado de forma reiterada significativamente superior al de las pruebas de criterio utilizadas en nuestra investigación.

En cuanto al tiempo y al esfuerzo que son necesarios para la obtención de esta información, considerada más valiosa para la intervención en función de las valoraciones realizadas por los profesores que participaron en la aplicación experimental del dispositivo, así como en función de los objetivos fundamentales de la evaluación dinámica (Sternberg y Grigorenko, 2003); ya hicimos mención en el capítulo de Instrumentos, cuando abordábamos el análisis del dispositivo EDPL, a la mayor relevancia de una información útil y eficaz para la optimización del proceso educativo de enseñanza y aprendizaje, con respecto al tiempo y al esfuerzo empleados. En cualquier caso, habría que distinguir entre una aplicación contextualizada del dispositivo y orientada a la intervención, que responda por tanto a nuestra propuesta

original, y la aplicación experimental que ha tenido lugar en nuestro estudio. En este sentido, ya recogimos en la exposición de nuestra propuesta original, así como en el análisis de la aplicación piloto del dispositivo, que una de las características fundamentales de EDPL es la *flexibilidad* de aplicación de las distintas actividades que lo componen y que, por tanto, el tiempo de aplicación estará en función de las necesidades de información sobre las dificultades de aprendizaje presentes, o sobre el modo óptimo de mejorar el proceso. El tiempo de aplicación se ve relacionado de esta forma con otro de los aspectos que consideramos cruciales en nuestra propuesta y que, así mismo, puede constituir un condicionante importante para su implementación en el contexto escolar; este es la *formación* necesaria para rentabilizar y optimizar el proceso de aplicación, y de esta forma poder extraer la información cualitativamente adicional y relevante.

Análisis de la información cualitativa recogida por los aplicadores del dispositivo EDPL

El análisis cualitativo del contenido de las Hojas de registro y evaluación

Ya hicimos mención con anterioridad a la recogida de información sobre el proceso de aplicación del dispositivo dinámico, efectuada por parte de los aplicadores. Esta información era anotada en las *Hojas de registro y evaluación* durante las sesiones de aplicación dinámica, y su contenido era fundamentalmente: (a) la evaluación del proceso de resolución de la actividad; (b) la mediación efectuada; (c) la evaluación de procesos metacognitivos; y por último, (d) observaciones. En las primeras sesiones del Grupo de Trabajo, previas a la implementación del tratamiento, así como durante el desarrollo del mismo, se instó a los profesores aplicadores a centrar su atención especialmente en el análisis de estos puntos.

Nuestro análisis de esta información se realizaría fundamentalmente atendiendo a los siguientes criterios:

- Valoración de los evaluadores participantes sobre la utilidad y eficacia de la aplicación dinámica del dispositivo. Incluye la valoración de la obtención de información valiosa para una posterior intervención, así como la valoración de un posible incremento en el rendimiento del alumnado, utilizando EDPL.
- Resultados obtenidos por el alumnado.
- Tiempo empleado (aunque esto venía dado en parte por el diseño de investigación).
- Tipo de información recogida y destacada por los evaluadores.
- Observaciones efectuadas.
- Patrones metodológicos de evaluación dinámica utilizados en la mediación.

Posteriormente, de estos seis criterios considerados inicialmente, el propio análisis del contenido de los protocolos nos hizo centrar nuestra atención especialmente en dos: *el tipo de información recogida y destacada por los aplicadores*, y los *patrones metodológicos de evaluación dinámica de la lectura utilizados en la mediación*. Las valoraciones de los aplicadores acerca del dispositivo EDPL no tuvieron lugar mayoritariamente en las narraciones recogidas en los protocolos, sino que éstas fueron efectuadas durante las sesiones del Grupo de Trabajo, así como durante las sesiones de evaluación conjunta final. Por su parte, los resultados obtenidos por el alumnado fueron

recogidos por los aplicadores en su valoración de cada uno de los indicadores de evaluación contemplados en la actividad, con lo que fueron posteriormente traducidos en puntuaciones dinámicas que han sido analizadas en páginas anteriores. Por otra parte, las valoraciones cualitativas sobre los resultados obtenidos por el alumnado han sido recogidas dentro del *tipo de información recogida y destacada por los aplicadores*. En cuanto al tiempo empleado en las sesiones de trabajo, pudimos ya analizar algunas de las dificultades destacadas por los aplicadores y provenientes en su mayoría de circunstancias relativas a la organización de la atención educativa en los centros o a situaciones de absentismo escolar. Por último, el criterio de observaciones ha sido incluido también en el tipo de información recogida.

El análisis cualitativo del contenido de las *Hojas de registro y evaluación* fue realizado mediante el establecimiento de un sistema de categorías. La categorización ha implicado la transformación y codificación de las observaciones, comentarios y análisis recogidos por los aplicadores, en *unidades de información significativas* para los objetivos de nuestro estudio (Lukas y Santiago, 2004). La selección e identificación de estas unidades de información en los protocolos entregados por los aplicadores, ha sido la base sobre la que se ha construido de forma progresiva nuestro sistema de categorías. En este sentido, aunque los objetivos del estudio estaban claros y, así mismo, el contenido de la información relevante que queríamos obtener a través del análisis, hemos partido de un sistema de categorías lo suficientemente abierto y flexible como para que la información revisada haya ido a su vez conformándolo y modificándolo continuamente. Por otra parte, la codificación de la información se ha realizado mediante la asignación de etiquetas verbales a las unidades de información significativas que fueron seleccionadas durante el desarrollo del análisis de contenido. A medida que iban siendo identificadas las categorías durante el desarrollo del análisis, éstas pasaban a formar parte del sistema, recibiendo *marcaciones* cuando eran nuevamente encontradas como fragmentos de información textual en los protocolos. La base textual sobre la que se ha realizado el análisis y la categorización, ha consistido en 1583 páginas extraídas de los 60 cuadernos de *Hojas de registro y evaluación*, así como de los anexos que utilizaron los aplicadores para informar sobre el proceso de aplicación de las actividades del dispositivo.

Tipo de información recogida y destacada por los aplicadores

Con relación al tipo de información recogida y destacada por los aplicadores, finalmente fueron identificadas 24 unidades de información significativas. Estas unidades de información funcionaron en realidad como supra-categorías, que aglutinaban a su vez observaciones y análisis recogidos por los aplicadores que podían tener entidad por sí mismas. La opción de agruparlas en grandes categorías respondió sencillamente a la necesidad de hacer más entendible y manejable la gran cantidad de información disponible. Por otra parte, el objetivo primordial de nuestro análisis acerca del tipo de información recogida, ha sido precisamente conocer y poder valorar el contenido de los datos anotados y destacados por los profesores que aplicaron el dispositivo; es decir, sobre qué aspectos o contenidos han centrado su atención finalmente, y sobre si el análisis de estas observaciones se relaciona con los resultados en cada uno de los grupos de aplicación. Este análisis se efectuaría mediante la valoración cualitativa del contenido y la cantidad de anotaciones en cada categoría recogida.

A continuación se recoge el sistema de categorías para el *tipo de información recogida y destacada por los aplicadores*, tal y como finalmente quedó elaborado. En éste podemos observar las 24 categorías, así como el tipo de anotaciones que fueron incluidas en las mismas. Tal y como exponíamos anteriormente, la construcción de este sistema de categorías ha seguido una doble vía; por una parte, algunas de las unidades de información significativas fueron identificadas de antemano, incluyéndose posteriormente las anotaciones del aplicador; por otra, también se ha operado de forma inductiva con la información anotada, ampliando y modificando las categorías y su contenido en función de la información destacada por los aplicadores. De este modo se han considerado al menos 57 unidades de información que finalmente fueron agrupadas en las 24 categorías contempladas.

TIPO DE INFORMACIÓN RECOGIDA Y DESTACADA POR LOS APLICADORES

1 La capacidad mostrada por el alumnado

Incluye anotaciones sobre:

- La capacidad mostrada por el alumno en la realización de una tarea determinada
- Las dificultades en la realización o puesta en práctica de habilidades

2 La reflexión en voz alta

Hace referencia a las reflexiones del alumno durante el desarrollo de las actividades de evaluación con relación a:

- Los contenidos de la tarea
- La puesta en práctica de determinadas estrategias
- Circunstancias relacionadas con aspectos de índole socio-personal
- Lo que los alumnos dicen de sí mismos con relación a la tarea

3 Los conocimientos del alumnado

Incluye anotaciones sobre los conocimientos que manifiesta un alumno en una tarea determinada

4 Tipo de dificultad relacionada con la lectura que presenta el alumno

Incluye anotaciones explícitas sobre el tipo de dificultad que presenta el alumno durante la realización de las actividades

5 La puesta en práctica de estrategias o habilidades de Asociación G-F

Incluye anotaciones sobre la puesta en práctica por parte del alumnado de los indicadores que sirven para valorar estos procesos

6 La puesta en práctica de estrategias o habilidades de coordinación y seguimiento visual de un texto

Incluye anotaciones sobre la puesta en práctica por parte del alumnado de los indicadores que sirven para valorar estos procesos

7 La puesta en práctica de habilidades de estructuración temporal

Incluye anotaciones sobre la puesta en práctica por parte del alumnado de los indicadores que sirven para valorar estos procesos

8 La puesta en práctica de estrategias de integración textual

Incluye anotaciones sobre la puesta en práctica por parte del alumnado de los indicadores que sirven para valorar estos procesos

9 La puesta en práctica de estrategias de integración texto-conocimientos y experiencias previas

Incluye anotaciones sobre la puesta en práctica por parte del alumnado de los indicadores que sirven para valorar estos procesos

10 La puesta en práctica de procesos de autorregulación y la presencia de metaconocimiento relacionado con la lectura

Incluye anotaciones sobre:

- La puesta en práctica por parte del alumnado de estrategias de planificación de la actividad, supervisión del proceso y autoevaluación
- La aplicación por parte del alumnado de estrategias a/de otras situaciones de aprendizaje
- La formulación o el reconocimiento por parte del alumnado de principios de carácter general, aplicables a distintas situaciones de aprendizaje relacionadas con la lectura
- Conciencia de las propias estrategias de comprensión, de cuándo, cómo y por qué son útiles
- Conciencia de lo que sabe acerca del contenido de los textos trabajados
- Conciencia sobre qué significa leer y comprender, y para qué sirve
- Conciencia de lo que sabe sobre la estructura de los textos
- Conciencia fonológica, léxica y sintáctica
- Conciencia sobre aspectos semánticos
- Conciencia de sus propias dificultades relacionadas con la tarea en particular o con la lectura en general
- Conciencia de sus propias actitudes, motivación e intereses relacionados con la lectura

11 Aspectos relacionados con los procesos de ajuste personal-social

Incluye anotaciones sobre:

- Cómo afronta el alumno o grupo de alumnos la tarea propuesta
- La actitud, motivación, interés mostrado con relación a la tarea o a la lectura en general
- Participación del alumnado en la actividad
- Autoestima, autoconcepto académico, tolerancia a la frustración
- Calidad de la colaboración con otros compañeros en el transcurso de la actividad

12 La puesta en práctica de estrategias de memoria y la atención mostrada

Incluye anotaciones sobre:

- La puesta en práctica de estrategias de memoria
- La utilización de los procesos de memoria a corto y largo plazo, así como de la memoria de trabajo
- El grado de atención y concentración mostrados por el alumnado durante la realización de la tarea
- La influencia de los procesos de memoria y atención en el desarrollo de la actividad del alumno

13 La puesta en práctica de estrategias intuitivas por parte del alumnado para realizar la actividad

Incluye anotaciones sobre las estrategias que utiliza el alumnado para acometer una actividad y que no forman parte de las estrategias contempladas en el dispositivo. Se trata habitualmente de estrategias de lectura *pasiva* (leer varias veces, repetir el texto para memorizarlo, copiar partes de un texto al realizar un resumen, etc.), o que son utilizadas como respuesta adaptativa a una dificultad propia, por ejemplo: leer mentalmente una palabra antes de decirla para asegurarse una buena decodificación.

14 Preguntas, respuestas, decisiones, explicaciones y propuestas del alumnado

15 La mejora del alumnado en la realización de la actividad durante el proceso mismo de mediación

Incluye anotaciones sobre:

- La respuesta positiva del alumnado a la mediación efectuada
- El incremento de rendimiento por parte del alumno o alumna, observado en el transcurso de una actividad o en el desarrollo de la aplicación del dispositivo

16 El proceso de mediación

Incluye anotaciones sobre:

- La actividad de mediación efectuada por parte del aplicador (preguntas realizadas para profundizar en un determinado contenido o para provocar la puesta en práctica de una estrategia, apoyo en la secuenciación de la actividad, explicación bidireccional para favorecer la toma de conciencia sobre algún proceso que se esté trabajando, introducir latencias, incrementar progresivamente el nivel de dificultad, simular situaciones que favorezcan un pensamiento estratégico, etc.)
- La valoración que efectúa el aplicador sobre la mediación efectuada
- La necesidad de mediación externa para acceder a los contenidos trabajados

17 Las circunstancias en que se desarrolla la aplicación del dispositivo

Incluye anotaciones sobre:

- Si las actividades han tenido lugar con un solo alumno o en grupo
- Ausencias del alumnado
- Dificultades en la aplicación derivadas del espacio físico donde se desarrolla, la falta de tiempo, la aplicación en grupo o la necesidad de compatibilizar la aplicación del dispositivo con la atención educativa a otros alumnos que no participan del estudio
- Las decisiones tomadas con relación a la puesta en práctica del dispositivo
- Número de sesiones empleadas en una actividad

18 Los procesos evaluados en la actividad y su valoración

Incluye anotaciones sobre:

- Los procesos implicados en la lectura que son objeto de estudio
- La valoración que el aplicador realiza de la contemplación de estos procesos para evaluar la comprensión
- Las posibilidades de evaluar estos procesos mediante el dispositivo EDPL
- Las actividades de evaluación propuestas

19 Los intentos del alumnado de realizar bien la tarea

20 La influencia de la mediación de los compañeros en la respuesta de un alumno

21 La realización de propuestas de intervención-evaluación por parte del aplicador

Incluye anotaciones sobre:

- Posibles *microtratamientos* para actuar sobre determinadas dificultades encontradas en la evaluación dinámica
- La necesidad de detenerse en alguno de los procesos evaluados para incidir en la mejora de un alumno
- La necesidad de ampliar el procedimiento de evaluación en alguna dirección determinada

22 La introducción y explicación de la actividad

23 Las dificultades en la realización de la mediación

24 La conducta, la actitud o el rendimiento del alumnado en otras situaciones de aprendizaje relacionadas con el proceso de evaluación dinámica que se lleva a cabo, o las circunstancias personales del alumno o alumna. Alusiones y valoración del aplicador que amplían o confirman la información recogida

Recogemos en la tabla que presentamos a continuación los valores medios obtenidos en cada una de las categorías contempladas. Estos valores nos hablarán de los aspectos que han centrado la atención de los aplicadores durante el transcurso de las sesiones de aplicación del dispositivo EDPL. Así mismo, en la tabla se recogen los niveles de correlación entre las anotaciones efectuadas para cada una de las categorías consideradas finalmente y la puntuación dinámica global obtenida en el dispositivo, y también la correlación del número de anotaciones con los valores anotados igualmente por los aplicadores relativos a la necesidad de mediación de los sujetos. Posteriormente analizaremos los datos más significativos de estas correlaciones, así como de las efectuadas entre las distintas categorías y las puntuaciones dinámicas relativas a cada uno de los procesos contemplados en el dispositivo de evaluación dinámica.

Tabla 85. Número de anotaciones efectuadas por los aplicadores para cada una de las categorías contempladas en función del tipo de información recogida y destacada en las Hojas de registro y evaluación (N = 35). Niveles de correlación del número de anotaciones en cada categoría con la puntuación dinámica global y con la necesidad de mediación.

Tipo de información recogida y destacada por los aplicadores	Suma	Media	Correlaciones	
			Punt. Dinámica global	Necesidad de mediación
Capacidad mostrada por el alumnado	694	19,83	0,068	0,166
Reflexión en voz alta	134	3,83	0,138	0,361*
Conocimientos del alumnado	4	0,11	-0,073	0,273
Tipo de dificultad relacionada con la lectura que presenta el alumno	1325	37,86	-0,608**	0,679**
Puesta en práctica de habilidades de Asociación G-F	195	5,57	0,335*	-0,067
Puesta en práctica de habilidades de Coordinación y seguimiento visual	38	1,09	0,098	0,086
Puesta en práctica de habilidades de Estructuración temporal	5	0,14	0,238	0,017
Puesta en práctica de estrategias de Integración textual	1184	33,83	-0,252	0,581**
Puesta en práctica de estrategias de Integración texto-conocimientos	209	5,97	-0,317	0,780**
Puesta en práctica de Procesos metacognitivos	1861	53,17	0,097	0,343*
Aspectos relacionados con los procesos de ajuste P-S	789	22,54	-0,027	0,494**
Puesta en práctica de estrategias de Memoria y Atención	492	14,06	-0,005	0,234
Puesta en práctica de Estrategias intuitivas	86	2,46	0,007	0,179
Preguntas, respuestas, decisiones, propuestas del alumnado	2165	61,86	-0,106	0,517**
Mejora durante el proceso de mediación	230	6,57	-0,267	0,446**
El proceso de mediación	1450	41,43	-0,381*	0,875**
Circunstancias en que se desarrolla la aplicación del dispositivo	151	4,31	0,435**	-0,189
Los procesos evaluados en EDPL y la valoración de los aplicadores	7	0,20	-0,182	0,410*
Los intentos del alumnado de realizar bien la tarea	26	0,74	0,033	0,185
Influencia de la mediación de los compañeros en las respuestas...	21	0,60	0,048	-0,059
Propuestas de evaluación-intervención de los aplicadores	1	0,02	0,066	0,167
Introducción y explicación de la actividad	99	2,83	0,278	-0,029
Dificultades en la realización de la mediación	1	0,02	0,066	0,167
Conducta, actitud o rendimiento en otras situaciones de aprendizaje	66	1,89	0,154	0,149
Total	11233	320,94		

Una primera inspección de los datos proporcionados en esta tabla, nos permite evidenciar el conjunto de anotaciones efectuadas sobre las unidades de información consideradas significativas para los objetivos del estudio. Por otra parte, también podemos ver el desigual reparto de estas anotaciones en función de las categorías contempladas. Este análisis nos permite de alguna forma acceder al contenido de las observaciones de los aplicadores y, de este modo, valorar de forma indirecta la relevancia que han otorgado a los distintos aspectos y procesos implicados durante el desarrollo de las sesiones de aplicación del dispositivo; por otra parte, nos permite valorar si la jerarquía establecida en cuanto a la relevancia otorgada en función de las observaciones realizadas, coincide básicamente con la propuesta realizada en el diseño de investigación, tanto desde el *Manual* entregado a los aplicadores, como desde las sesiones de formación en el Grupo de Trabajo, previas a la implementación del dispositivo.

En este sentido, podemos verificar que, tanto los procesos metacognitivos, como el proceso de mediación, han ocupado un lugar muy relevante entre los aspectos sobre los que han centrado su atención los aplicadores. Concretamente, tras las observaciones relativas a las respuestas otorgadas por los sujetos experimentales a las cuestiones planteadas durante la aplicación del dispositivo, los procesos metacognitivos y el proceso mismo de mediación llevada a cabo, han sido las categorías que recibieron un mayor número de anotaciones. Este dato confirmaría el hecho de que los aplicadores han centrado de forma especial su atención en aquellos aspectos considerados *clave* desde la propuesta inicial, y que estaban recogidos en las *Hojas de registro y evaluación*. Ligeramente por detrás en el número de anotaciones, se encuentran el tipo de dificultad concreta que presentaba el alumnado participante y las observaciones relacionadas con la puesta en práctica de los procesos de integración textual. El resto de procesos implicados en la lectura, y contemplados en el dispositivo, han recibido también la atención de los aplicadores, aunque ésta ha sido desigual; de esta forma, los procesos de ajuste personal-social han recibido una atención considerable a juzgar por el número de anotaciones relativas a éstos que efectuaron los aplicadores. Con un número significativamente menor de anotaciones, se encuentran los procesos de integración texto-conocimientos y los procesos de asociación grafema-fonema. A pesar de que el número de indicadores de los procesos de integración texto-conocimientos, era inicialmente superior al de integración textual o al de los procesos de ajuste, recibieron una menor atención por parte de los aplicadores a tenor del número de anotaciones realizadas; probablemente, el mayor número de actividades realizadas pertenecientes al bloque de procesos de integración textual, facilitó que los aplicadores centraran su atención, y también sus observaciones, en estos procesos. Por otra parte, en cuanto a los procesos de ajuste, el anterior razonamiento no sería válido para explicar el mayor número de observaciones realizadas, ya que ni siquiera se llegó a realizar la actividad específica propuesta para la evaluación de estos procesos. La relevancia otorgada a estos procesos desde la propuesta realizada, plasmada en la presencia de indicadores de evaluación en la casi totalidad de tareas, y puesta de relieve también por los aplicadores, incluyendo hasta en 14 ocasiones indicadores de estos procesos en la evaluación de los sujetos cuando las tareas propuestas no los contenían, parece haberse reflejado en la cantidad de observaciones relativas a estos procesos efectuadas por los aplicadores. Por último, en cuanto a los procesos psicológicos subyacentes, éstos han visto repartidas las observaciones efectuadas, ya que, por una parte, algunas anotaciones fueron realizadas específicamente sobre habilidades de coordinación y seguimiento visual, o sobre habilidades de estructuración temporal; estas anotaciones son muy escasas y dejan

entrever que los aplicadores, incluso en tareas de tipo no lingüístico, centraron su atención prioritariamente en aspectos relativos a procesos de autorregulación y de ajuste personal-social. Por otra parte, sí recibieron una atención especial los procesos de memoria y atención. Dentro de estos, debemos destacar las observaciones relativas a la utilización de estrategias de memoria, así como al mantenimiento de la atención durante el desarrollo de las actividades.

Otros aspectos que llaman nuestra atención, son los relativos a la mejora de los sujetos durante el proceso de mediación, con una considerable cantidad de observaciones; así como aquellos relacionados con la capacidad mostrada por los sujetos durante el desarrollo de las actividades. El primero de los datos, sobre todo por cuanto no era algo sobre lo que fuera solicitada información en el estudio, ofrece información relevante acerca de la constatación del aplicador, en el transcurso del proceso, de que las capacidades del sujeto experimentaban una cierta mejora. Por otra parte, las circunstancias relativas a la aplicación del dispositivo, también han sido objeto de observaciones por parte de los aplicadores. En este sentido, aunque la categoría de dificultades en la realización de la mediación aparece con una sola anotación, la relativa a las circunstancias en las que se desarrolló la aplicación, recoge observaciones que, de forma indirecta, aluden a posibles problemas que dificultaron, cuando no impidieron, una aplicación ideal de la propuesta recogida en el Manual. Estas circunstancias, como podemos comprobar en el contenido de las categorías elaboradas, han recogido todos aquellos problemas a los que hemos hecho alusión en diferentes momentos de nuestro informe, relativos a las ausencias del alumnado participante, las dificultades derivadas de la aplicación grupal del dispositivo o las derivadas de la necesidad de llevar a cabo el tratamiento experimental mientras eran atendidos otros alumnos no participantes en el estudio. Por último, una ampliación de los datos que aparecen en la tabla, nos lleva a constatar que, tal y como ocurre con la mayoría de las categorías elaboradas, el número de anotaciones fue muy desigual entre los grupos de aplicación. Precisamente este hecho nos llevó a hipotetizar una posible relación entre el simple número de anotaciones efectuadas, totalmente o en función de una de las categorías, y los resultados obtenidos por un grupo de sujetos.

A continuación exponemos los datos más significativos de la totalidad de los análisis correlacionales efectuados. Debemos hacer mención en primer lugar al elevado número de correlaciones significativas entre las 24 categorías contempladas en el sistema elaborado. En este sentido, el análisis de fiabilidad de las anotaciones efectuadas nos muestra un valor de correlación inter-ítem de $r_{0,39}$, con una gran mayoría de correlaciones significativas o muy significativas estadísticamente. El valor del coeficiente *Alfa* fue también bastante alto, situándose en 0,87. Esto nos hablaría positivamente acerca de una considerable consistencia interna del sistema de categorías construido.

El primer dato de interés que nos muestra la tabla expuesta anteriormente es la correlación negativa y muy significativa entre el *tipo de dificultad específica* y la puntuación dinámica global; como tendremos oportunidad de comprobar para la práctica totalidad de los procesos contemplados en EDPL, las anotaciones relativas al tipo de dificultad se han relacionado de forma clara con la presencia de mayores dificultades en el alumnado y, como consecuencia, con un menor nivel de puntuación otorgada en función del proceso de aprendizaje. En línea con lo expresado anteriormente, el contraste con el nivel de correlación observado entre el *tipo de*

dificultad específica y la *necesidad de mediación*, hace patente los mayores requerimientos de apoyo del aplicador, por parte de los sujetos que presentaron más dificultades en el desarrollo de las actividades.

Por otra parte, las anotaciones sobre la *puesta en práctica de habilidades de asociación G-F* correlacionan positivamente con la puntuación dinámica global. Esto no ocurrirá, sin embargo, con ninguna otra categoría alusiva a la puesta en práctica de los distintos procesos; de este modo, el número de anotaciones relativas a la puesta en práctica de un proceso, no se habría correspondido finalmente con la obtención de altas puntuaciones dinámicas. En efecto, en la valoración del profesorado realizada en función de los indicadores de evaluación propuestos, fue valorada, lógicamente, la *calidad* de su puesta en práctica y no sólo la *presencia* de ese indicador, que es lo que sí reflejan las anotaciones recogidas en nuestro sistema de categorías.

Son destacables, por otra parte, los niveles de correlación obtenidos entre la puesta en práctica de los procesos de *integración textual*, *integración texto-conocimientos*, *procesos metacognitivos* y *procesos de ajuste P-S*, y la *necesidad de mediación*. Esto nos hablaría claramente de que las anotaciones sobre la puesta en práctica de estos procesos, no siempre han destacado una buena ejecución del sujeto; antes bien, estos datos parecen ser indicativos de que la mayoría de las anotaciones han reflejado una puesta en práctica no exenta de dificultades por parte de los sujetos. La verificación de lo que hemos expresado se realiza al comprobar que la correlación entre la necesidad de mediación y la puntuación dinámica obtenida en los distintos procesos contemplados en el dispositivo, es negativa y significativa. Por otra parte, observamos que el nivel de correlación entre la *mejora durante el proceso de mediación* y la puntuación dinámica global, en contra de lo que inicialmente cabría esperar, es negativo. Posteriormente, en el capítulo de Discusión, abordaremos con mayor profundidad las razones que podrían explicar este resultado.

En cuanto a las anotaciones sobre el *proceso de mediación*, éstas han sido claramente relativas a la necesidad de mediación por parte del sujeto; ya hemos visto que este dato se ha correspondido con la obtención de puntuaciones dinámicas más bajas. En este sentido podemos observar que la correlación con la *necesidad de mediación* es altísima, quedando prácticamente identificados ambos valores. Por otra parte, las anotaciones efectuadas sobre las *circunstancias en que se desarrolla la aplicación*, correlacionaron positiva y significativamente con la puntuación dinámica global. A pesar de que el contenido de la información incluida en esta categoría, hacía esencialmente referencia a aspectos relativos a circunstancias problemáticas, que habrían tenido una incidencia negativa en el desarrollo de las sesiones de aplicación, y, como consecuencia, en los resultados obtenidos, tenemos que constatar que aquellos grupos de aplicación en los que se realizaron mayor cantidad de anotaciones sobre estas circunstancias, obtuvieron mejores puntuaciones dinámicas.

Por último, destacar las correlaciones positivas, aunque sin llegar a la significatividad estadística, que mantiene la categoría *introducción y explicación de la actividad*, con la puntuación dinámica global; así como la que mantiene con esta misma puntuación el número de anotaciones sobre la *actitud, conducta o rendimiento en otras situaciones de aprendizaje*. Este último dato sería indicativo de que, mayoritariamente, las referencias a la puesta en práctica de actitudes o conductas relacionadas con el desarrollo de las

actividades, en otras situaciones de aprendizaje, se habrían correspondido con la obtención de mayores puntuaciones dinámicas.

Relación entre el tipo de información recogida y destacada por los aplicadores, y las puntuaciones dinámicas

En cuanto a los datos más significativos de las correlaciones entre el tipo de información destacada por los aplicadores y las puntuaciones dinámicas obtenidas por los sujetos en los distintos procesos contemplados en EDPL, destacamos a continuación éstos en función del proceso contemplado en el dispositivo.

En primer lugar, con relación a la puntuación dinámica obtenida en los procesos psicológicos subyacentes, ésta mantiene una correlación negativa y muy significativa con la anotación relativa al *tipo de dificultad específica de lectura* ($r.-0,525$). Este dato, que como veremos se repite para la práctica totalidad de procesos considerados en EDPL, parece lógico si pensamos que un mayor número de anotaciones sobre la presencia de dificultades específicas en la comprensión habría tenido su reflejo en la obtención de más bajas puntuaciones en el dispositivo. También es negativa y significativa la correlación con la *puesta en práctica de los procesos de integración texto-conocimientos* ($r.-0,391$), así como con las anotaciones sobre el *proceso de mediación efectuado* ($r.-0,472$). En este último caso, resultado que también se repite para otros procesos, parece haber regido la lógica de que a mayor necesidad de mediación, puntuaciones finales más bajas. Por otra parte, con relación a las *circunstancias en que se desarrolló la aplicación*, la correlación es positiva y significativa ($r.0,387$). Igualmente correlaciona positivamente, aunque no de forma significativa, con la *puesta en práctica de habilidades de coordinación y seguimiento visual* ($r.0,017$) y con las *habilidades de estructuración temporal* ($r.0,159$).

En cuanto a los procesos de asociación G-F, éstos correlacionan negativamente y de forma muy significativa con las anotaciones sobre el *tipo de dificultad específica de lectura* ($r.-0,472$), así como también con la *puesta en práctica de procesos de integración texto-conocimientos* ($r.-0,334$). Con el *proceso de mediación*, al igual que ocurría en los procesos psicológicos subyacentes, la correlación es igualmente negativa y muy significativa ($r.-0,466$). Por otra parte, la correlación con las *circunstancias en que se desarrolla la aplicación* es positiva, aunque carece de significatividad estadística ($r.0,111$), al igual que con la *puesta en práctica de estrategias de asociación G-F*, siendo en este caso el nivel de correlación muy bajo ($r.0,093$).

Para la puntuación dinámica de los procesos de integración textual, la correlación con el *tipo de dificultad específica* es también negativa y significativa ($r.-0,414$). Con el *proceso de mediación* correlaciona negativamente, aunque en este caso el nivel es más bajo que en anteriores procesos ($r.-0,270$). Por último, con las anotaciones sobre las *circunstancias en que se desarrolla la aplicación*, la correlación vuelve a ser positiva y significativa ($r.0,343$).

Con relación a la puntuación en los procesos de integración texto-conocimientos previos, hay que decir que el comportamiento de las correlaciones ha sido distinto al observado en el resto de procesos. En primer lugar, aparece con un nivel de correlación positiva y significativa la *reflexión en voz alta de los sujetos* ($r.0,368$), y no correlaciona negativamente con el *tipo de dificultad específica*. Por otra parte, la correlación que

mantiene esta puntuación con las anotaciones sobre la *puesta en práctica de los procesos de integración texto-conocimientos* sí es significativa ($r.0,356$). Mantiene igualmente unos niveles de correlación positivos, aunque no significativos, con la *puesta en práctica de habilidades de estructuración temporal* ($r.0,263$), con la *puesta en práctica de estrategias de integración textual* ($r.0,205$), de *estrategias metacognitivas* ($r.0,199$), de *procesos de ajuste personal-social* ($r.0,263$), y con la *puesta en práctica de estrategias intuitivas* ($r.0,194$). También mantiene una correlación positiva, y significativa en este caso, con *los intentos del sujeto de realizar bien la tarea propuesta* ($r.0,372$). En cuanto al *proceso de mediación*, la correlación, al contrario de lo que ha ocurrido en el resto de procesos, es positiva ($r.0,284$), y, sin embargo, negativa para las *circunstancias en que se desarrolló la aplicación* ($r.-0,161$).

La puntuación dinámica de los procesos metacognitivos vuelve a correlacionar negativa y muy significativamente con el *tipo de dificultad específico* ($r.-0,574$). Mantiene un nivel de correlación positivo con la *puesta en práctica de los procesos de asociación G-F* ($r.0,308$), así como con la *puesta en práctica de los procesos metacognitivos* ($r.0,108$). Por su parte, también correlaciona negativamente con las anotaciones sobre el *proceso de mediación* ($r.-0,320$) y positivamente, llegando a la significatividad estadística, con las anotaciones relativas a las *circunstancias en que se desarrolló la aplicación del dispositivo* ($r.0,409$).

La puntuación en los procesos de ajuste también correlaciona negativa y muy significativamente con el *tipo de dificultad específica de lectura* anotada por los aplicadores ($r.-0,451$). En cambio, mantiene correlaciones positivas con la *puesta en práctica de procesos de asociación G-F* ($r.0,333$), *habilidades de coordinación y seguimiento visual* ($r.0,228$), de *estructuración temporal* ($r.0,285$), de *estrategias metacognitivas* ($r.0,225$), y de los propios *procesos de ajuste personal-social* ($r.0,093$). Los niveles de correlación han sido positivos con la puesta en práctica de todos los procesos contemplados a excepción de los *procesos de asociación G-F*. Por otra parte, también mantiene una correlación positiva, y muy significativa en este caso, con las anotaciones sobre las *circunstancias de aplicación* ($r.0,467$). Así mismo, mantiene niveles medios de correlación positiva con varias categorías que no aparecen con esos niveles en los contrastes anteriores. Por ejemplo, con *la influencia de la mediación de los compañeros* ($r.0,274$); con *los intentos del alumno por realizar bien la tarea* ($r.0,174$); con *la introducción y la explicación de la actividad* ($r.0,311$), con las *dificultades en la realización de la mediación* ($r.0,211$), y, finalmente, con las anotaciones relativas a *la actitud y la conducta del sujeto en otras situaciones de aprendizaje* ($r.0,248$). Por último, con el *proceso de mediación*, la correlación es negativa, aunque no significativa ($r.-0,249$).

En síntesis, el análisis de las correlaciones de las puntuaciones dinámicas del dispositivo con las distintas categorías contempladas, nos ofrece también información acerca del contenido real de las anotaciones efectuadas por los aplicadores. En este sentido, hemos podido comprobar reiteradamente que las anotaciones sobre el proceso de mediación han consistido mayoritariamente en la constatación de la necesidad de mediación por parte del sujeto, lo que hemos visto verificado también mediante la alta correlación positiva entre este dato y el que hacía referencia al número de veces que el sujeto necesitó la mediación. Por otra parte, también es destacable que los niveles de correlación entre las distintas puntuaciones dinámicas y las anotaciones sobre la puesta en práctica de los procesos correspondientes, no hayan sido muy altos. Aunque esta

correlación ha sido positiva en todos los casos, tan sólo para los procesos de integración texto-conocimientos ha presentado significatividad. La explicación de estos bajos niveles de correlación puede estar sencillamente en que el número de anotaciones sobre la puesta en práctica de los distintos procesos, no se ha visto siempre reflejado linealmente en la valoración que posteriormente los aplicadores han realizado en las *Hojas de registro y evaluación*. Esta valoración incluía lógicamente la calidad de esa puesta en práctica, por lo que los niveles de correlación se habrían visto afectados.

Patrones metodológicos de evaluación dinámica de la lectura utilizados en la mediación

El sistema de categorías construido con relación a los patrones metodológicos de evaluación dinámica utilizados por los aplicadores en el proceso de mediación, se estructuró finalmente en tres grandes apartados: (a) *aspectos generales*, que contemplaba básicamente patrones de mediación de tipo general, no ligados específicamente a un bloque de procesos determinado; (b) *mediación del conocimiento metacognitivo y del uso de estrategias*, en el que se recogieron todos aquellos patrones de mediación relativos a los procesos metacognitivos —incluyendo el conocimiento metacognitivo y las estrategias de autorregulación, transferencia y generalización—; los procesos de integración textual y los procesos de integración del texto con los conocimientos y experiencias previas, todos ellos contemplados en el dispositivo EDPL; y, por último, (c) *mediación de procesos de ajuste personal-social*, donde quedaron recogidos todos aquellos patrones de mediación relativos a los procesos de desarrollo socio-emocional contemplados también en nuestra propuesta. En el marco de estos tres grandes apartados fueron consideradas 19 grandes categorías, que aglutinaron a su vez 178 unidades de información significativas para los objetivos de nuestro estudio. A diferencia de la consideración que tuvieron las unidades de información en el sistema de categorías elaborado para el tipo de información recogida y destacada por los aplicadores, en el que las anotaciones realizadas para distintas unidades de información eran sumadas, anotadas e integradas finalmente como una sola categoría; en el sistema construido para los patrones metodológicos de evaluación dinámica, las unidades de información contempladas fueron consideradas de forma autónoma en la matriz de datos construida, otorgándose a cada una el número de anotaciones correspondiente a los fragmentos de texto relativos a éstas. Esta opción respondía mejor al objetivo de identificar de la forma más precisa posible aquellos patrones metodológicos de evaluación dinámica que habrían resultado claves en la explicación de las mejoras observadas en los sujetos, como, así mismo, en la explicación de las mejoras observadas en el proceso mismo de evaluación, valoradas por los profesores aplicadores. De esta forma, hemos podido «desgranar» mejor la información relativa a los patrones metodológicos utilizados en la mediación, pudiendo acceder con mayor precisión a la estructura del proceso de mediación puesto en práctica durante la aplicación experimental del dispositivo.

A continuación exponemos, en primer lugar, el sistema de categorías elaborado para los *patrones metodológicos de evaluación dinámica de la lectura utilizados en la mediación*. La construcción progresiva de este sistema de categorías ha seguido un proceso similar al descrito para el tipo de información recogida por los aplicadores. Al igual que entonces, la construcción ha seguido una doble vía, caracterizada en primera instancia por la identificación previa de algunas de las unidades de información

significativas, entre las cuales se encontraban, lógicamente, los patrones metodológicos de evaluación dinámica propuestos ya en el Manual entregado a los aplicadores al comienzo del estudio. Las anotaciones de los profesores relativas a la puesta en práctica de conductas mediadoras ya recogidas de antemano, fueron incluyéndose en estas unidades de información, conformando a su vez las categorías coincidentes con los procesos contemplados en la propuesta. Por otra parte, a medida que las *Hojas de registro y evaluación* que contenían la información anotada por los aplicadores, iban siendo analizadas minuciosamente, también se han ido incluyendo, a la vez que ampliando o modificando, las unidades de información inicialmente consideradas, e incluso las propias categorías y su contenido. La ampliación significativa, así como la estructuración del sistema de categorías construido finalmente, ha respondido sin duda al trabajo desarrollado por el profesorado que realizó la aplicación experimental de EDPL. El volumen de información y la relevancia de la mayoría de las anotaciones y observaciones efectuadas, nos ha permitido analizar y extraer un conjunto de patrones metodológicos que podrían resultar significativos para la obtención de información útil con relación a tres aspectos: (a) la mejora del proceso de aprendizaje del sujeto; (b) la mejora del proceso de enseñanza y mediación del profesorado; (c) una mejor y mayor comprensión de dónde se encuentran las dificultades de lectura y de cómo optimizar el proceso de lectura y comprensión.

La clave identificativa que precede a cada una de las unidades de información contempladas, ha tenido una función fundamentalmente práctica, y está compuesta de: una letra minúscula, que indica el orden en el que aparece la categoría; tres letras iniciales, que responden al contenido mismo de la categoría; y, por último, un número, que muestra la secuencia de las distintas unidades de información.

Posteriormente, mostramos una tabla donde se recoge de forma resumida el volumen de anotaciones realizadas por los aplicadores para cada una de las unidades de información contempladas, así como el número de anotaciones medias para cada una de las categorías construidas (tabla 86). Seguidamente, exponemos en la tabla 87 los resultados de los análisis correlacionales efectuados entre las anotaciones realizadas en cada uno de los patrones metodológicos de mediación y las distintas puntuaciones dinámicas. Hemos considerado a este efecto las puntuaciones del dispositivo EDPL, tanto la puntuación global como las relativas a los distintos procesos contemplados, así como las puntuaciones de mejora de las pruebas criterio: ECO 1, ECOS, APSL y CATTELL. También hemos incluido en esta tabla la variable *Necesitó mediación*, por considerar de interés las relaciones que pueda mostrar con los distintos patrones de mediación, sobre todo en función de los resultados que hemos venido mostrando hasta ahora relacionados con esta variable. Por último, mostramos la relación entre el número de anotaciones efectuadas en cada uno de los grupos de aplicación y la puntuación dinámica global. La determinación de posibles diferencias en la utilización de los patrones de mediación y su posible correlación con los resultados obtenidos, es a priori uno de los objetivos relevantes de este análisis. Mostraremos aquí las relaciones que mantienen ambos elementos a nivel cuantitativo; más tarde, cuando abordemos el análisis cualitativo de la aplicación efectuada por los aplicadores, tendremos oportunidad de completar y matizar algunos de los efectos obtenidos.

PATRONES METODOLÓGICOS DE EVALUACIÓN DINÁMICA DE LA LECTURA UTILIZADOS EN LA MEDIACIÓN

Aspectos generales

1(a) Profundización en la respuesta del alumno/a o en su justificación, haciendo nuevas preguntas o cuestionando sus respuestas

- a.pro1 *Profundizar en la respuesta del alumno/a o en su justificación, haciendo nuevas preguntas o cuestionando sus respuestas*

2(b) Ampliación de la actividad mediante la realización de tareas relacionadas que surgen del propio alumnado o del evaluador

- b.amp2 *Ampliar la actividad con la intención de practicar las estrategias trabajadas*
b.amp3 *Ampliar la actividad con la intención de obtener información adicional o complementaria*

3(c) Establecimiento de apoyo en la comprensión, secuenciación o realización de la actividad

- c.est4 *Facilitar la comprensión de la actividad (explicar en qué consiste, poner algún ejemplo, utilizar la ayuda de dibujos, modificar una cuestión para facilitar su comprensión por parte del alumnado, etc.)*
c.est5 *Introducir la actividad para favorecer una mayor comprensión de la misma*
c.est6 *Acompañar en el desarrollo y realización de la actividad*
c.est7 *Establecer apoyo gradualmente mediante preguntas hasta poder llegar a inducir las respuestas correctas*
c.est8 *Incrementar el nivel de dificultad*
c.est9 *Ceder gradualmente el control de la actividad al alumno*
c.est10 *Explicar o aclarar un concepto con relación a la demanda de un alumno o por ser necesario durante el desarrollo de la actividad*
c.est11 *Explicar por qué una actividad no puede ser realizada de una determinada forma*
c.est12 *Secuenciar el contenido de la actividad para facilitar su realización*
c.est13 *Lanzar cuestiones al grupo de alumnos/as para provocar su atención, reflexión y respuesta*

4(d) Recapitulación conjunta del trabajo realizado anteriormente con el objetivo de retomarlo

- d.rec14 *Reconstruir colectivamente lo trabajado*
d.rec15 *Preguntar por actividades realizadas anteriormente*

5(e) Mediación de la discusión y la reflexión sobre contenidos, elementos, estrategias y procesos

- e.dis16 *Fomentar la elaboración conjunta de estrategias, hipótesis, significados o conclusiones, mediante el lanzamiento de cuestiones, la explicación o introducción intercalada de algún concepto o la puesta de relieve de alguna idea expresada por un/a alumno/a*
e.dis17 *Discutir y valorar conjuntamente el contenido -literal o inferencial- de un texto, ampliando información, exponiendo dudas, etc.*

e.dis18 *Favorecer mediante explicación bidireccional la toma de conciencia sobre los procesos, elementos o estrategias que se están trabajando*

6(f) Repetición de una actividad para poder poner en práctica las estrategias trabajadas y para poder evaluar su incorporación a la ejecución de la actividad por parte del alumno

f.rep19 *Pedir al alumno que vuelva a leer, poniendo en práctica las estrategias aprendidas*

f.rep20 *Volver a realizar una actividad con el objetivo de que el alumno pueda poner en práctica estrategias trabajadas y para poder valorar su incorporación a la ejecución de la actividad por parte del alumno*

f.rep21 *Realizar una segunda vez la actividad sin que haya existido mediación explícita por parte del evaluador*

Mediación del conocimiento metacognitivo y del uso de estrategias

Mediación del conocimiento metacognitivo relacionado con la lectura

• **7(g) Conciencia de lo que sabe sobre elementos estructurales y tipologías de textos**

g.est22 *Preguntar al alumno para evaluar su conciencia estructural: ¿Sabrías decirme a qué tipo de textos pertenece esta lectura? ¿Cómo lo sabes?*

g.est23 *Preguntar: ¿Por qué crees que este texto está desordenado? ¿Sabrías ordenarlo tú?*

g.est24 *Preguntar si la estructura en la que se presenta un texto es correcta y porqué*

g.est25 *Preguntar si sabe dónde se coloca habitualmente el título de un texto o cualquier otro elemento estructural que el evaluador considere oportuno*

g.est26 *Preguntar al alumno/a: ¿Cuál es el título de este texto?*

g.est27 *Preguntar: ¿Para qué sirve el título en un texto? ¿Qué nos dice? ¿Es importante? ¿Por qué?*

g.est28 *Leer junto al alumno varios párrafos de un texto desordenado previamente y pedirle que elija el que iría al comienzo del texto*

g.est29 *Favorecer o pedir explícitamente la comparación entre diversos textos para evaluar su estructura*

g.est30 *Preguntar ante la lectura de un texto previamente desordenado: ¿Ocurre algo extraño en la historia que has leído?*

g.est31 *Favorecer la comparación entre textos ordenados y desordenados para facilitar la toma de conciencia sobre la mayor comprensión de los textos ordenados y de sus claves estructurales*

g.est32 *Preguntar: ¿Podríamos colocar un título en este texto sin leerlo? ¿Por qué?*

g.est33 *Preguntar: ¿Cómo comienza habitualmente un cuento o una historia? ¿Cómo termina?*

g.est34 *Favorecer la valoración del alumno acerca de si diversos textos presentados están ordenados o desordenados y por qué*

• **8(h) Conciencia sobre las propias estrategias de comprensión, de cuándo, cómo y por qué son útiles**

h.con35 *Hablar de las estrategias que hemos utilizado*

h.con36 *Favorecer mediante explicación bidireccional la toma de conciencia sobre las estrategias que podemos utilizar en una determinada actividad*

h.con37 *Hablar de las estrategias que son necesarias para poder realizar de forma óptima una tarea determinada*

- h.con38 *Preguntar al final de la actividad: ¿Qué debemos hacer para realizar esta tarea de la mejor forma posible? ¿Qué estrategias deberíamos poner en práctica?*
- h.con39 *Preguntar al alumno: ¿Qué estrategias has utilizado para realizar la actividad? ¿Cómo la has hecho?*
- h.con40 *Preguntar durante la realización de la tarea: ¿Qué hacéis para poder realizar la actividad? ¿Qué estrategias estáis poniendo en práctica?*
- h.con41 *Preguntar al final: ¿Qué es necesario para poder hacer la actividad?*
- h.con42 *Preguntar al alumno/a: ¿Qué podríamos hacer cuando finalizamos una lectura para mejorar nuestra comprensión? ¿Qué estrategias podemos utilizar?*
- h.con43 *Preguntar: ¿Qué es necesario saber para poder realizar bien esta tarea?*
- h.con44 *Preguntar: ¿Qué podríamos hacer para comprender mejor?*
- h.con45 *Preguntar al alumno si conoce otras estrategias para realizar la actividad u otra forma de hacerla*
- h.con46 *Preguntar al final de una actividad: ¿Piensas que has realizado esta actividad con la suficiente atención?*
- h.con47 *Preguntar: ¿Has tenido que estar especialmente atento para hacer esta actividad? ¿Por qué?*
- h.con48 *Preguntar al alumno: Si crees que la actividad no está bien, ¿Qué podrías hacer?*
- h.con49 *Preguntar durante la lectura: ¿Cómo empiezas la actividad? ¿Qué estrategias pones en práctica?*
- h.con50 *Preguntar: ¿Qué has hecho cuando te has equivocado durante la lectura?*
- h.con51 *Preguntar: ¿Cómo podemos saber lo que tenemos que hacer en una actividad?*
- h.con52 *Preguntar: ¿Crees que es importante conocer el significado de las palabras para comprender un texto? ¿Y si no las conocemos? ¿Qué podríamos hacer?*
- h.con53 *Preguntar: ¿Cómo sabemos que está bien la actividad que hemos hecho?*
- h.con54 *Preguntar: ¿Cómo podemos saber de qué va un texto sin leerlo?*
- h.con55 *Ante un texto que se presenta incompleto, se pregunta al alumno si podría continuar la historia sin haberla leído; se le cuestiona acerca de qué sería necesario para poder realizar bien la actividad*

- **9(i) Conciencia de lo que sabe acerca del contenido de los textos trabajados o de una actividad determinada**

- i.sab56 *Preguntar si ha comprendido*
- i.sab57 *Preguntar a un/a alumno/a sobre su conocimiento acerca de lo que sabe sobre un determinado contenido, o sobre si cree haber comprendido una actividad o un texto, o realizado bien una actividad*
- i.sab58 *Preguntar a un/a alumno/a: ¿Sabrías contar lo que has leído? o ¿Sabrías realizar esta actividad?*
- i.sab59 *Favorecer mediante explicación bidireccional la toma de conciencia sobre cómo podemos saber que hemos comprendido, o preguntar: ¿Cómo podemos saber que hemos comprendido bien un texto?*
- i.sab60 *Preguntar: ¿Necesitarías volver a leer la lectura para comprender mejor?*

- **10(j) Conciencia de la estructura fonológica, léxica y sintáctica del lenguaje**

- j.fl61 *Preguntar al alumno si conoce a qué categoría de palabras pertenecen las que estamos trabajando en ese momento*
- j.fl62 *Preguntar para qué cree que sirven los signos de puntuación*
- j.fl63 *Preguntar: Cuando estás leyendo... ¿Tienes en cuenta los signos de puntuación? ¿Cómo lo haces?*
- j.fl64 *Preguntar: ¿Son importantes los signos de puntuación? ¿Por qué?*

- j.fl65 *Preguntar: Cuando terminamos de escribir un texto... ¿Qué ponemos?*
- j.fl66 *Preguntar en el transcurso de una actividad relacionada con la estructuración sintáctica: ¿Por qué has colocado esta palabra en este lugar?*
- j.fl67 *Favorecer o pedir explícitamente la comparación entre diversos textos para establecer diferencias en función de la presencia o no de signos de puntuación*
- j.fl68 *Pedir al alumno que valore la corrección de un texto: ¿Está bien escrito el texto que has leído? ¿Le ocurre algo? ¿Le falta algo?...*
- j.fl69 *Preguntar al alumno si conoce el nombre de la parte de la oración que ha estado manejando en el transcurso de una actividad relacionada con la estructuración sintáctica*
- j.fl70 *Preguntar: ¿Dónde hay que hacer pausas cuando leemos? ¿Por qué?*
- j.fl71 *Pedir al alumno/a que coloque los signos de puntuación a un texto*
- j.fl72 *Preguntar en el transcurso de una actividad relacionada con la estructuración sintáctica: ¿Qué tipo de palabra vendría a continuación de esta que has colocado?*
- j.fl73 *Mantener y transcribir conversaciones para favorecer la toma de conciencia acerca de las palabras funcionales y su uso*

- **11(k) Conciencia de las propias dificultades relacionadas con la lectura**

- k.dif74 *Analizar conjuntamente la importancia de reconocer las propias dificultades*
- k.dif75 *Preguntar al alumno: ¿Por qué crees que realizamos esta actividad?*

Mediación del uso de estrategias de lectura

12(l) Mediación del uso de estrategias de planificación

- l.pla76 *Preguntar al alumno antes de comenzar, qué cree que hay que hacer en la tarea propuesta*
- l.pla77 *Preguntar al alumno: ¿Qué debemos hacer para realizar esta tarea de forma adecuada? ¿Qué estrategias sería necesario poner en práctica?*
- l.pla78 *Preguntar al grupo de alumnos/as: ¿Qué haríais para enseñar a un/a compañero/a vuestro/a cómo se hace esta actividad de forma que la tuviera muy bien?*
- l.pla79 *Simular una situación en la que el alumno debe explicar a un/a compañero/a el contenido de una actividad o cómo solucionarla*
- l.pla80 *Preguntar: ¿Qué es necesario para poder hacer la actividad?*
- l.pla81 *Preguntar por el objetivo o el propósito con el cual leemos un determinado texto o realizamos una actividad*
- l.pla82 *Establecer apoyo en la identificación del objetivo de la tarea*
- l.pla83 *Preguntar en el transcurso de la actividad: ¿Qué es necesario hacer a partir de ahora?*
- l.pla84 *Preguntar: ¿Cómo vas a hacer la actividad?*
- l.pla85 *Preguntar: ¿Qué es lo primero que hay que hacer en esta actividad? ¿Y después?...*
- l.pla86 *Construir colectivamente una secuencia de operaciones -a modo de estrategia- a realizar para afrontar la actividad y hacerla bien*
- l.pla87 *Preguntar: Si comenzáramos de nuevo la actividad... ¿Cómo la realizarías?*

13(m) Mediación del uso de estrategias de supervisión

- m.sup88 *Pedir al alumno que realice anotaciones sobre el texto que está leyendo. Estas anotaciones pueden estar relacionadas con la dificultad que supone la lectura de ese texto, con su posible procedencia o tipología, con las dudas o*

cuestiones que le surjan al leerlo, con conocimientos o experiencias previas que le evoquen, etc.

- m.sup89 *Tratar de focalizar la atención del alumno/a sobre la comprensión del texto que se está leyendo o sobre algún proceso trabajado*
- m.sup90 *Realizar una segunda vez la actividad focalizando la atención sobre un aspecto determinado del proceso de lectura*
- m.sup91 *Pedirle explícitamente que supervise o revise sus respuestas*
- m.sup92 *Pedirle explícitamente que revise su actividad con el objetivo de encontrar posibles errores*
- m.sup93 *Analizar junto al grupo de alumnos/as un texto (corregir o revisar elementos léxicos, sintácticos o semánticos)*
- m.sup94 *Trasladar a los alumnos la supervisión y evaluación de la actividad en general o del trabajo realizado por un compañero en particular*
- m.sup95 *Introducir latencias, tratando de provocar una lectura más reflexiva y favoreciendo un mayor control y una mejor regulación del proceso de comprensión*
- m.sup96 *Establecer apoyo en la elaboración y puesta en práctica de estrategias de memoria*
- m.sup97 *Pedir al alumno que subraye lo más importante del texto o aquello que no entiende*
- m.sup98 *Leer en voz alta al alumno para que éste descubra posibles absurdos e incoherencias*
- m.sup99 *Preguntar durante la tarea: ¿Qué es lo que tenías que hacer? o ¿Recuerdas lo que tenías que hacer en esta actividad?*
- m.sup100 *Apoyar verbalmente la reconstrucción del objetivo de la tarea y de la secuencia de operaciones a realizar*
- m.sup101 *Apoyar a un alumno a encontrar el lugar por el que iba desarrollando su lectura o su actividad: Recuerda y fíjate bien en lo último que hiciste y continúa por ahí....*
- m.sup102 *Pedir al alumno que lea en voz alta para favorecer la detección de posibles errores*
- m.sup103 *Leer en voz alta al alumno para modelar el uso de estrategias de supervisión de la propia comprensión*
- m.sup104 *Apoyar la construcción (p.e. el acceso al significado de una palabra) del significado a través del contexto*
- m.sup105 *Establecer apoyo en la identificación de sonidos que componen las palabras y en su escritura, mediante la focalización de la atención en el proceso, el uso de la mediación verbal y de la autoevaluación*
- m.sup106 *Establecer apoyo mediante preguntas para que el alumno descubra absurdos e incoherencias. ¿Encuentras algo extraño en estas palabras? ¿Las conoces?...*

14(n) Mediación del uso de estrategias de autoevaluación

- n.aut107 *Preguntar por la dificultad o facilidad de una determinada tarea de lectura para provocar la puesta en práctica de estrategias de autoevaluación de la propia comprensión*
- n.aut108 *Preguntar: ¿Cuál ha sido más difícil para ti? ¿por qué?*
- n.aut109 *Preguntar si una determinada actividad es fácil o difícil y porqué*
- n.aut110 *Preguntar: ¿En qué consiste la dificultad de esta tarea?*
- n.aut111 *Lanzar cuestiones al grupo de alumnos/as para que evalúen conjuntamente la dificultad de una actividad y detecten dónde y porqué han cometido errores*
- n.aut112 *Lanzar cuestiones al grupo de alumnos/as para que analicen conjuntamente los mecanismos que les han llevado a comprender y realizar de forma adecuada una determinada tarea*

- n.aut113 *Preguntar en el transcurso o al final de la actividad: ¿Qué es lo que teníamos que hacer en esta actividad? ¿Para qué la estamos haciendo? ¿Para qué nos ha servido?*
- n.aut114 *Preguntar al final de la actividad: ¿Qué hemos hecho? ¿Qué hemos aprendido?*
- n.aut115 *Preguntar al grupo de alumnos/as sobre los elementos de los que consta la actividad que hemos realizado*
- n.aut116 *Favorecer o pedir explícitamente la comparación entre dos textos o actividades para evaluar su dificultad*
- n.aut117 *Preguntar al grupo de alumnos/as qué estrategia de las que hemos usado es más adecuada para poder realizar la actividad*
- n.aut118 *Preguntar al grupo de alumnos/as: Si alguien no supiera nada de la lectura que hemos leído y leyera este resumen...¿Entendería bien de qué va el texto?*
- n.aut119 *Preguntar: ¿Cómo crees que has realizado la actividad?*
- n.aut120 *Favorecer la valoración del alumno acerca de las razones que le han llevado a mejorar en su actividad*
- n.aut121 *Recordarle o pedirle explícitamente que evalúe su propia actividad una vez finalizada*
- n.aut122 *Preguntar: ¿Te ha gustado esta lectura? ¿Por qué?*
- n.aut123 *Preguntar: ¿Cuál de los dos textos que has leído te ha gustado más? ¿Por qué?*
- n.aut124 *Preguntar: ¿Habéis comprobado que está bien vuestra actividad? ¿Estáis seguros/as?*
- n.aut125 *Recordar la importancia de evaluar la propia actividad*
- n.aut126 *Apoyar mediante la realización de preguntas la autoevaluación del alumno, o ir analizando la actividad por partes (ir corrigiendo junto a él paso a paso)*
- n.aut127 *Preguntar: ¿Has finalizado tu actividad? ¿Está terminada?*
- n.aut128 *Preguntar: ¿Tú crees que tiene sentido esta oración? ¿Por qué?*
- n.aut129 *Pedirle que explique lo que ha leído con la intención de que evalúe su comprensión*
- n.aut130 *Preguntar: Mientras ibas leyendo... ¿Estabas pensando en lo que leías?*
- n.aut131 *Preguntar: ¿Qué texto has leído mejor? ¿Cuál has comprendido mejor? ¿Por qué?*

15(ñ) Mediación del uso de estrategias de transferencia y generalización

- ñ.tra132 *Poner en relación la actividad que estamos realizando con otras situaciones de aprendizaje en las que sea útil poner en práctica las estrategias trabajadas o de las que sea beneficioso aplicar estrategias a la actividad actual*
- ñ.tra133 *Preguntar al alumno al finalizar una actividad: ¿Qué hemos aprendido? ¿Para qué nos sirve?*
- ñ.tra134 *Preguntar al alumno sobre la utilidad de aprender una estrategia o habilidad determinada: ¿Puede servirnos lo que estamos trabajando para comprender mejor? ¿Por qué?*
- ñ.tra135 *Preguntar: ¿Puede servirnos lo que hemos hecho para alguna otra actividad? ¿Para cuál?*

16(o) Mediación del uso de estrategias de integración textual

- o.tex136 *Hablar de los distintos tipos de textos que estamos trabajando y de las claves estructurales que facilitan la comprensión*
- o.tex137 *Establecer apoyo en la explicación de una respuesta o una experiencia relacionada con la lectura, en la construcción de un concepto o en la justificación de una respuesta*
- o.tex138 *Pedir a un alumno que exponga su respuesta, la explique o la justifique*

- o.tex139 *Pedir a un alumno que intente exponer, completar o justificar su respuesta de otra forma cuando sus compañeros/as no han logrado entenderle o cuando el evaluador lo considera oportuno*
- o.tex140 *Solicitar del alumno que cuente de qué va la lectura que ha leído*
- o.tex141 *Preguntar antes de comenzar qué tipo de lectura (tipología de texto) vamos a leer*
- o.tex142 *Solicitar del alumno que lea un texto para contar posteriormente de qué va o para realizar un resumen del mismo*
- o.tex143 *Pedir al alumno que lea en voz alta o en silencio para favorecer la concentración o la detección de un determinado elemento*
- o.tex144 *Establecer apoyo gradual para favorecer la comprensión y el uso de una estrategia estructural*
- o.tex145 *Hacer preguntas sobre el contenido de la lectura, o relacionadas con el tema, para favorecer la integración y la representación mental del contenido*
- o.tex146 *Modelar el uso de los signos de puntuación en la lectura de un texto, para favorecer la observación y la toma de conciencia por parte del alumno/a.*
- o.tex147 *Favorecer mediante explicación bidireccional el uso de estrategias de categorización para definir o explicar el significado de las palabras*
- o.tex148 *Práctica guiada de una estrategia para favorecer la representación mental del contenido de un texto (ir haciendo preguntas y comentarios acerca de ilustraciones que se corresponden o no con el texto trabajado)*
- o.tex149 *Establecer apoyo en la identificación de ideas relevantes y de la síntesis global de un texto, así como en la elaboración de un resumen por parte del alumno*
- o.tex150 *Favorecer el establecimiento de relaciones entre el contenido de un texto y el título*
- o.tex151 *Establecer apoyo en la comprensión de un texto mediante la secuenciación de la historia y la realización de dibujos por parte del alumno*
- o.tex152 *Establecer apoyo en la búsqueda de información en el texto*
- o.tex153 *Solicitar al alumno que coloque un título que sea adecuado para un texto o que juzgue si un determinado título es adecuado y por qué*
- o.tex154 *Facilitar la estructuración sintáctica mediante preguntas dirigidas a las distintas partes de la oración o señalando las distintas partes con colores*
- o.tex155 *Establecer apoyo en el acceso al significado de palabras mediante la presentación de ilustraciones*

17(p) Mediación del uso de estrategias para la activación de conocimientos previos y la integración de conocimientos y experiencias con nuevos contenidos

- p.con156 *Preguntar a los alumnos si tienen conocimientos o experiencias relacionadas con el contenido que aparece en el texto que se está trabajando, o pedirles que imaginen la situación descrita*
- p.con157 *Favorecer mediante explicación bidireccional la toma de conciencia sobre la relación que existe entre acceder a nuestros conocimientos y experiencias previas, y la mayor comprensión de un texto*
- p.con158 *Poner en relación la actividad que estamos realizando con experiencias y conocimientos previos*
- p.con159 *Establecer apoyos para facilitar el establecimiento de relaciones entre diversos textos o diversas actividades*
- p.con160 *Preguntar al grupo de alumnos/as, antes de comenzar a leer un texto, sobre qué creen que trata, para favorecer la formulación de hipótesis y la utilización de elementos estructurales para comprender*
- p.con161 *Preguntar al grupo de alumnos qué creen que pasará en la lectura que van a leer*

- p.con162 *Preguntar una vez iniciada la lectura: ¿Qué crees que pasará ahora? ¿Qué crees que pasará al final de la lectura?*
- p.con163 *Preguntar en el transcurso de la lectura o al final: ¿Qué hubiera pasado si...?*
- p.con164 *Preguntar al alumno una vez que ha leído: Mientras ibas leyendo... ¿Has pensado en lo que podría pasar al final? ¿Y en lo que tú dijiste que podría pasar?*
- p.con165 *Preguntar al final de la lectura: ¿Qué dijiste tú que pasaría? ¿Ha pasado lo que tú dijiste? ¿Cómo termina la historia de verdad?*
- p.con166 *Favorecer la reflexión sobre la relación que existe entre formular hipótesis durante la lectura y la mejor comprensión del texto*
- p.con167 *Establecer apoyo para facilitar la activación de conocimientos previos*
- p.con168 *Pedir al alumno que imagine la situación descrita en un texto y relate qué haría él en esa situación*

Mediación de procesos de ajuste personal-social

18(q) Mediación de la motivación, la actitud y el interés con relación a la lectura

- q.mot169 *Tratar de reconducir a un alumno que abandona el objetivo de la tarea por motivos relacionados con la actitud o la motivación hacia la lectura*
- q.mot170 *Establecer apoyo de tipo afectivo-emocional a un alumno con relación a su capacidad o a su trabajo*
- q.mot171 *Implicar al alumnado en la preparación de la actividad o en la toma de decisiones acerca del desarrollo de misma*
- q.mot172 *Proponer la actividad en formato de juego o enigma a resolver*
- q.mot173 *Reforzar positivamente o confirmar la respuesta o la actitud del alumno/a*

19(r) Mediación de la participación y el apoyo entre los componentes del grupo

- r.par174 *Facilitar el apoyo entre compañeros/as del grupo en la explicación, resolución o valoración de una actividad*
- r.par175 *Facilitar la participación de todos/as en las actividades de discusión, valoración y puesta en común*
- r.par176 *Hacer intervenir a algún compañero para que muestre o guíe en la búsqueda de otras posibles respuestas, o para que exponga conocimientos o experiencias relacionadas con la actividad*
- r.par177 *Hacer alusión a una respuesta dada por uno de los alumnos, destacándola como modelo de estrategia o razonamiento válido para realizar mejor la actividad*
- r.par178 *Facilitar la puesta en común de anotaciones, explicaciones y trabajo realizado, para ser discutido y valorado conjuntamente por el grupo de alumnos/as*

Tabla 86. Número de anotaciones para cada una de las unidades de mediación

Mediación	Suma	Media	Mediación	Suma	Media	Mediación	Suma	Media
a.pro1	497	14,20	i.sab59	3	0,09	n.aut120	1	0,03
1(a) total	497	14,20	i.sab60	2	0,06	n.aut121	13	0,37
b.amp2	5	0,14	9 (i) total	176	5,03	n.aut122	11	0,31
b.amp3	100	2,86	j.fl61	8	0,23	n.aut123	1	0,03
2 (b) total	105	3,00	j.fl62	12	0,34	n.aut124	12	0,34
c.est4	103	2,94	j.fl63	2	0,06	n.aut125	2	0,06
c.est5	5	0,14	j.fl64	2	0,06	n.aut126	10	0,29
c.est6	78	2,23	j.fl65	1	0,03	n.aut127	5	0,14
c.est7	136	3,89	j.fl66	3	0,09	n.aut128	6	0,17
c.est8	12	0,34	j.fl67	29	0,83	n.aut129	1	0,03
c.est9	6	0,17	j.fl68	11	0,31	n.aut130	3	0,09
c.est10	40	1,14	j.fl69	1	0,03	n.aut131	1	0,03
c.est11	2	0,06	j.fl70	3	0,09	14 (n) total	268	7,66
c.est12	16	0,46	j.fl71	4	0,11	ñ.tra132	27	0,77
c.est13	21	0,60	j.fl72	1	0,03	ñ.tra133	2	0,06
3 (c) total	419	11,97	j.fl73	1	0,03	ñ.tra134	7	0,20
d.rec14	8	0,23	10 (j) total	78	2,23	ñ.tra135	1	0,03
d.rec15	4	0,11	k.dif74	2	0,06	15 (ñ) total	37	1,06
4 (d) total	12	0,34	k.dif75	2	0,06	o.tex136	2	0,06
e.dis16	27	0,77	11 (k) total	4	0,11	o.tex137	23	0,66
e.dis17	4	0,11	l.pla76	39	1,11	o.tex138	28	0,80
e.dis18	25	0,71	l.pla77	3	0,09	o.tex139	4	0,11
5 (e) total	56	1,60	l.pla78	8	0,23	o.tex140	108	3,09
f.rep19	21	0,60	l.pla79	2	0,06	o.tex141	2	0,06
f.rep20	23	0,66	l.pla80	0	0,00	o.tex142	8	0,23
f.rep21	13	0,37	l.pla81	1	0,03	o.tex143	3	0,09
6 (f) total	57	1,63	l.pla82	6	0,17	o.tex144	5	0,14
g.est22	4	0,11	l.pla83	4	0,11	o.tex145	93	2,66
g.est23	3	0,09	l.pla84	12	0,34	o.tex146	4	0,11
g.est24	13	0,37	l.pla85	3	0,09	o.tex147	6	0,17
g.est25	7	0,20	l.pla86	7	0,20	o.tex148	2	0,06
g.est26	9	0,26	l.pla87	3	0,09	o.tex149	3	0,09
g.est27	12	0,34	12 (l) total	88	2,51	o.tex150	1	0,03
g.est28	4	0,11	m.sup88	2	0,06	o.tex151	1	0,03
g.est29	5	0,14	m.sup89	87	2,49	o.tex152	2	0,06
g.est30	19	0,54	m.sup90	12	0,34	o.tex153	8	0,23
g.est31	10	0,29	m.sup91	45	1,29	o.tex154	22	0,63
g.est32	7	0,20	m.sup92	3	0,09	o.tex155	1	0,03
g.est33	4	0,11	m.sup93	12	0,34	16 (o) total	326	9,31
g.est34	2	0,06	m.sup94	20	0,57	p.con156	13	0,37
7 (g) total	99	2,83	m.sup95	24	0,69	p.con157	7	0,20
h.con35	9	0,26	m.sup96	14	0,40	p.con158	12	0,34
h.con36	15	0,43	m.sup97	3	0,09	p.con159	13	0,37
h.con37	13	0,37	m.sup98	12	0,34	p.con160	13	0,37
h.con38	23	0,66	m.sup99	7	0,20	p.con161	6	0,17
h.con39	42	1,20	m.sup100	5	0,14	p.con162	9	0,26
h.con40	4	0,11	m.sup101	1	0,03	p.con163	15	0,43
h.con41	1	0,03	m.sup102	2	0,06	p.con164	10	0,29
h.con42	2	0,06	m.sup103	2	0,06	p.con165	8	0,23
h.con43	2	0,06	m.sup104	1	0,03	p.con166	1	0,03
h.con44	8	0,23	m.sup105	3	0,09	p.con167	4	0,11
h.con45	6	0,17	m.sup106	11	0,31	p.con168	1	0,03
h.con46	1	0,03	13 (m) total	266	7,60	17 (p) total	112	3,20
h.con47	1	0,03	n.aut107	46	1,31	q.mot169	7	0,20
h.con48	13	0,37	n.aut108	33	0,94	q.mot170	12	0,34
h.con49	2	0,06	n.aut109	41	1,17	q.mot171	14	0,40
h.con50	2	0,06	n.aut110	3	0,09	q.mot172	13	0,37
h.con51	1	0,03	n.aut111	11	0,31	q.mot173	28	0,80
h.con52	2	0,06	n.aut112	5	0,14	18 (q) total	74	2,11
h.con53	6	0,17	n.aut113	18	0,51	r.par174	18	0,51
h.con54	1	0,03	n.aut114	12	0,34	r.par175	17	0,49
h.con55	1	0,03	n.aut115	2	0,06	r.par176	3	0,09
8 (h) total	155	4,43	n.aut116	22	0,63	r.par177	12	0,34
i.sab56	82	2,34	n.aut117	3	0,09	r.par178	9	0,26
i.sab57	84	2,40	n.aut118	2	0,06	19 (r) total	59	1,69
i.sab58	5	0,14	n.aut119	4	0,11	Σ total	2888	82,52

Tabla 87. Valores de correlación entre los patrones de mediación utilizados y las puntuaciones dinámicas

<i>Puntuaciones dinámicas</i> <i>Patrones de mediación</i>	Procesos P. Subyacentes	Asociación G-F	Integración textual	Integración texto-con.	Procesos metacog.	Procesos de ajuste	Puntuación global	Necesitó mediación ¹	Mejora ECO 1	Mejora ECOS	Mejora APSL	Mejora CI
1(a) Profundización en la respuesta	-0,345*	-0,428*	-0,160	0,214	-0,174	-0,158	-0,244	0,691**	0,133	-0,734*	-0,068	0,349*
2 (b) Ampliación de la actividad	-0,069	-0,214	-0,056	0,174	-0,001	0,140	-0,008	0,232	0,158	-0,001	-0,012	0,082
3 (c) Apoyo en la comprensión...	-0,315	-0,453**	-0,269	0,258	-0,228	-0,132	-0,278	0,618**	-0,189	-0,784**	0,134	0,039
4 (d) Recapitulación conjunta...	-0,211	-0,455**	-0,108	0,255	0,067	0,100	-0,004	0,372*	0,219	-0,505	-0,027	0,299
5 (e) Mediación de la discusión ...	0,186	-0,057	0,136	0,211	0,118	0,281	0,148	0,064	0,207	-0,756**	-0,058	0,024
6 (f) Repetición de una actividad...	-0,381*	-0,572**	-0,276	0,128	-0,209	-0,174	-0,283	0,600**	0,050	-0,593	0,114	0,310
7 (g) Conciencia sobre la estructura...	-0,609**	-0,464**	-0,295	0,083	-0,416**	-0,446**	-0,492**	0,779**	-0,021	-0,080	-0,227	0,320
8 (h) Conciencia sobre las propias est...	-0,166	-0,192	-0,093	0,276	-0,047	0,086	-0,065	0,502**	0,190	-0,600	-0,159	0,194
9 (i) Conciencia de lo que sabe...	-0,320	-0,489**	-0,225	0,302	-0,161	-0,130	-0,241	0,732**	0,146	-0,745**	-0,154	0,248
10 (j) Conciencia de la estructura fon...	-0,550**	-0,556**	-0,335*	0,060	-0,426*	-0,457**	-0,511**	0,753**	-0,151	0,403	0,146	0,150
11 (k) Conciencia de las propias dif.	-0,109	-0,012	-0,123	-0,001	-0,133	0,011	-0,134	0,209	-0,065	0,153	-0,247	0,054
12 (l) Mediación de est. de planif.	-0,302	-0,362*	-0,152	0,290	-0,105	-0,023	-0,160	0,664**	0,214	-0,601	-0,182	0,313
13 (m) Mediación de est. de superv.	-0,331	-0,467**	-0,301	0,222	-0,275	-0,162	-0,319	0,513**	-0,135	-0,796**	0,293	0,045
14 (n) Mediación de est. de autoev.	-0,140	-0,375*	-0,081	0,364*	-0,032	0,096	-0,060	0,482**	0,217	-0,802**	0,010	0,131
15 (ñ) Mediación de est. de transf..	-0,007	-0,198	-0,073	0,264	0,053	0,207	0,049	0,223	0,163	-0,576	0,019	0,023
16 (o) Mediación de est. de int. textual	-0,764**	-0,682**	-0,427*	0,140	-0,548**	-0,584**	-0,637**	0,825**	-0,052	-0,565	-0,064	0,237
17 (p) Mediación de est. de int. tex-con.	-0,382*	-0,269	-0,202	0,306	-0,257	-0,227	-0,293	0,708**	-0,093	-0,534	-0,041	0,224
18 (q) Mediación de la motivación...	0,037	-0,174	-0,004	0,341*	0,172	0,266	0,155	0,243	0,063	-0,580	0,127	0,017
19 (r) Mediación de la participación...	0,061	0,049	0,046	0,157	-0,008	0,164	0,037	0,042	0,167	-0,386	-0,078	-0,023

¹ Esta variable no constituye obviamente una de las puntuaciones dinámicas; su inclusión en esta tabla obedece al interés que revelan los valores de correlación establecidos con los patrones de mediación.

La revisión de los datos extraídos de los análisis descriptivos y correlacionales efectuados, nos muestra información relevante que merece ser valorada detenidamente. En primer lugar, se desprende de los datos que la mayoría de los patrones de mediación habrían actuado como indicadores indirectos de la existencia de mayores dificultades en los sujetos experimentales, tal y como ya ocurriera con algunos de los tipos de información recogida y destacada por los aplicadores. Esto se plasma en los valores de correlación negativa que muestran muchos patrones de mediación con las distintas puntuaciones dinámicas obtenidas mediante la aplicación del dispositivo EDPL. Otros patrones metodológicos, en cambio, habrían sido indicativos de mediación conducente de forma general a la obtención de puntuaciones dinámicas altas en el dispositivo.

La matriz de correlaciones entre los 19 patrones de mediación contemplados, muestra una mayoría de correlaciones positivas y muy significativas entre ellos. Este resultado expresa por una parte la estrecha relación existente entre algunos de estos patrones metodológicos, de forma que la puesta en práctica de algunos de ellos ha supuesto a su vez la puesta en práctica de otros. El análisis de estas correlaciones muestra un coeficiente de correlación inter-ítems de 0,45 para los 19 patrones de mediación y de 0,14 para las 178 unidades de información. Por otra parte, el coeficiente de fiabilidad (Alfa de Cronbach) para los 178 elementos de la matriz es de 0,92, mientras que para los 19 patrones de mediación es de 0,91.

Como podemos comprobar, la mayoría de las correlaciones establecidas entre los distintos patrones metodológicos de mediación y las puntuaciones dinámicas del dispositivo EDPL, son negativas. Esto parece indicar de forma clara que un mayor número de anotaciones por parte de los aplicadores, o lo que es lo mismo, la mayor puesta en práctica de determinadas conductas mediadoras, ha supuesto la presencia de mayores dificultades relacionadas con los procesos que se estaban evaluando. Esto ha ocurrido de forma especial en algunos procesos. En este sentido, parece claro con relación a la puntuación dinámica obtenida en los procesos de Asociación G-F, que una mayor «intervención» por parte de los aplicadores, ha sido indicativa de la presencia de mayores problemas y, por tanto, de la obtención de más bajas puntuaciones. Como podemos observar, la práctica totalidad de los contrastes muestran valores de correlación negativos, y un número importante de ellos presentan correlaciones significativas estadísticamente. Sin embargo, como tendremos oportunidad de comprobar posteriormente, es importante destacar que la constatación de estas correlaciones negativas entre patrones de mediación y puntuaciones dinámicas EDPL, no significa necesariamente que los sujetos que necesitaron de mediación, aunque obtuvieran bajas puntuaciones, no modificaran en sentido positivo sus actuaciones entre las fases pretest y posttest como consecuencia de la intervención realizada. Por otra parte, la observación de los datos también nos muestra que el número de correlaciones negativas, así como la fuerza de estas correlaciones, disminuye a medida que valoramos las puntuaciones dinámicas de aquellos procesos que resultaron más difíciles para los sujetos.

Como apuntábamos anteriormente, las correlaciones negativas entre patrones y puntuaciones dinámicas no significarían necesariamente una ausencia de mejora postratamiento de los sujetos experimentales. La constatación de este hecho podemos encontrarla en los valores de correlación obtenidos para las puntuaciones de mejora. En efecto, los valores observados ya no son tan negativos en su mayoría, y la fuerza de las correlaciones negativas es mucho menor. Estos resultados son generalizados salvo para

las puntuaciones de mejora de la prueba ECOS, donde la puesta en práctica de la mayor parte de las conductas de mediación no se ha relacionado con la obtención de mejoras postest; antes bien, se ha relacionado negativamente con éstas. Paradójicamente, los sujetos experimentales de educación secundaria obtuvieron mejoras muy significativas en la prueba de comprensión, tal y como hemos analizado en el capítulo de resultados en la batería criterio. Una primera conclusión de estos resultados apuntaría hacia la no relevancia del proceso de mediación en las mejoras observadas para este grupo. La puntuación postest obtenida en esta prueba por los sujetos de secundaria, correlacionaba positiva y significativamente con la mejora, pero no así con la puntuación dinámica EDPL, con la que, tanto el postest como la mejora correlacionaban negativamente. Aunque la revisión de las puntuaciones dinámicas obtenidas por los sujetos de secundaria en esta prueba, revela la importancia de éstas, el análisis de regresión que efectuábamos y mostrábamos cuando acometíamos el análisis del valor predictivo para las distintas pruebas, indicaba claramente que la puntuación dinámica global no participaba de la explicación de los resultados en la puntuación de mejora. Sólo dos patrones metodológicos de mediación: *conciencia de la estructura fonológica, léxica y sintáctica*, y *conciencia de las propias dificultades*, mantienen una correlación positiva con la puntuación de mejora.

Sin embargo, por lo que respecta a las puntuaciones de mejora obtenidas en la prueba de comprensión de primaria, observamos que la mayoría de los valores de correlación son positivos. Revisando los análisis sobre el valor predictivo de las puntuaciones de mejora en esta prueba, comprobamos que la puntuación dinámica contribuía a explicar de forma significativa las mejoras observadas. Este dato confirmaría la relevancia que habría tenido la mediación ejercida por los aplicadores, mediación que ahora vemos concretada en la puesta en práctica de los distintos patrones metodológicos. Aunque sin llegar a la significatividad estadística, algunos de estos patrones parecen haber tenido una relevancia notable en la explicación de los resultados; tales son: *la recapitulación conjunta del trabajo realizado*, *la mediación de la discusión y reflexión sobre las estrategias de comprensión*, *la mediación de la conciencia sobre las propias estrategias de comprensión*, *la mediación del uso de estrategias metacognitivas de planificación y autoevaluación*, etc.

Por su parte, la puntuación de mejora obtenida en la escala APSL muestra un mayor número de valores de correlación negativos, aunque con escasa fuerza. El *apoyo en la comprensión*, *la mediación del uso de estrategias de supervisión*, y *la mediación de la motivación, la actitud y el interés*, serían los patrones metodológicos que mayor relación guardarían con las mejoras observadas.

Por último, las correlaciones entre los patrones de mediación y la puntuación de mejora en la prueba de inteligencia, denotan la relevancia que habría tenido la puesta en práctica de la mediación en la explicación de las mejoras postratamiento. En efecto, la revisión del análisis de regresión correspondiente también nos informa de que la puntuación dinámica global explicaba de forma significativa la mejora en la puntuación CI. Observamos en este caso que la totalidad de los patrones de mediación correlacionan positivamente con la mejora, llegando uno de ellos a la significatividad estadística. *La profundización en la respuesta del alumno o en su justificación, haciendo nuevas preguntas o cuestionando sus respuestas*; *la recapitulación conjunta del trabajo realizado*, *la repetición de una actividad con el objetivo de evaluar la puesta en práctica de un determinado proceso o de ampliar la información*; *la mediación de la*

conciencia sobre elementos estructurales, sobre las propias estrategias de comprensión, así como sobre lo que el sujeto sabe acerca del contenido de los textos trabajados; la mediación del uso de estrategias de planificación, de estrategias de integración textual o de estrategias de integración del texto con conocimientos previos; son algunos de los patrones de mediación que mejor se relacionan con la mejora observada en la prueba de inteligencia.

En función de estos resultados, hemos querido comprobar cómo se comportaban los datos cuando seleccionábamos a los sujetos que obtuvieron mejoras posttest. En este sentido, hemos observado que, en general, cuando se seleccionaron los sujetos *mejoradores*, los niveles de correlación experimentaron incrementos notables; estos resultados se dieron, sobre todo, con relación a los procesos de ajuste P-S y a los procesos metacognitivos, aunque sin llegar a mostrar finalmente niveles significativos. En la relación entre los patrones de mediación de estrategias de transferencia y la puntuación dinámica de los procesos metacognitivos, obtenida para aquellos sujetos que obtuvieron mejoras en la prueba ECO 1 partiendo de bajas puntuaciones, observamos un incremento de los valores de correlación que va desde $r.0,053$ para el conjunto de la muestra a $r.0,217$ para los mejoradores. También se deja notar este incremento con relación a la puntuación dinámica global, con un descenso de correlaciones negativas entre los distintos patrones y la puntuación obtenida. En general, con relación a la mejora en ECO 1 ($N = 9$), se produce un incremento notable en los valores de correlación, aunque ningún valor llega a la significatividad.

Por su parte, cuando se seleccionaron los sujetos, o grupos, que obtuvieron mejoras en la prueba ECOS partiendo de bajas puntuaciones iniciales ($N = 11$), obtenemos valores de correlación positivos y muy significativos entre algunos patrones de mediación y la puntuación dinámica de los procesos de integración texto-conocimientos. Estos patrones son los siguientes:

3 (c) Apoyo en la comprensión...	$r. 0,947^{**}$
4 (d) Recapitulación conjunta...	$r.0,847^{**}$
6 (f) Repetición de una actividad...	$r.0,775^{**}$
9 (i) Conciencia de lo que sabe...	$r.0,669^*$
13 (m) Mediación de est. de superv.	$r.0,893^{**}$
14 (n) Mediación de est. de autoev.	$r.0,827^{**}$
15 (ñ) Mediación de est. de transf..	$r.0,976^{**}$
18 (q) Mediación de la motivación...	$r.0,963^{**}$

También hay un incremento notable de los valores de correlación con relación a los procesos metacognitivos y a los procesos de ajuste para esta prueba, aunque sin llegar a la significación estadística. Por otra parte, similar a lo expuesto para los procesos de integración texto-conocimientos en la prueba de comprensión de secundaria, es lo que obtuvimos cuando seleccionamos los sujetos que obtuvieron ganancias en la escala APSL ($N = 13$). En este caso, los resultados de los contrastes indicaron valores de correlación significativos para los patrones:

1(a) Profundización en la respuesta	r.0,562*
8 (h) Conciencia sobre las propias est...	r.0,645*
9 (i) Conciencia de lo que sabe...	r.0,589*
14 (n) Mediación de est. de autoev.	r.0,588*

Por último, cuando se seleccionaron los sujetos mejoradores en la prueba de inteligencia (N = 25), también observamos incremento en los valores de correlación de algunos patrones de mediación con la puntuación dinámica de los procesos de integración texto-conocimientos, pasando algunas de estas a ser significativas estadísticamente.

3 (c) Apoyo en la comprensión...	r.0,417*
9 (i) Conciencia de lo que sabe...	r.0,404*
14 (n) Mediación de est. de autoev.	r.0,452*
18 (q) Mediación de la motivación...	r.0,462*

Para estos mejoradores en las puntuaciones CI, la correlación que mantienen los patrones de mediación *1(a) Profundización en la respuesta...* y *16 (o) Mediación de estrategias de int. textual*, con la puntuación postest de ECO 1, resulta significativa (r.0,604 para el primer contraste y r.0,675 para el segundo). Por otra parte, la mediación de los procesos de integración textual correlaciona positiva y significativamente con la mejora en CI (r.0,411). En general, hay un incremento muy notable de los valores de correlación para los mejoradores con respecto a esta prueba.

A continuación mostramos en la tabla 88 la relación entre los patrones metodológicos de evaluación dinámica y la puntuación dinámica global, en función de los grupos de aplicación contemplados en el estudio. Los datos expuestos muestran, al igual que ya lo hicieran los que hacían referencia al grupo experimental en su conjunto, una escasa relación entre el número de anotaciones efectuadas con relación a los patrones de mediación, y la puntuación dinámica obtenida; antes bien, esta relación es mayoritariamente negativa e incluso significativa en algunos casos. De cualquier modo, los datos mostrados son necesariamente limitados, ya que los cálculos para el coeficiente de correlación entre los patrones de mediación y la puntuación dinámica para los grupos que realizaron las anotaciones de forma grupal, no pueden ser calculados. La simple inspección de los datos, sin embargo, no parece mostrar un cambio en los valores de correlación, por cuanto no se observa de forma clara ninguna tendencia en los resultados. Por otra parte, la determinación de las diferencias en la aplicación del dispositivo en función del grupo de aplicación, sobre todo en lo concerniente a la mediación efectuada, era uno de nuestros objetivos de análisis. En este sentido, las diferencias en cuanto al número de anotaciones efectuadas, son claramente observables. Un análisis de varianza muestra diferencias globales estadísticamente muy significativas entre los distintos grupos ($F = 66,539$; $p < 0,000$). Igualmente, cuando se realiza este contraste para cada uno de los patrones de mediación, las diferencias son muy significativas en todos los casos, excepto en *11(K) Conciencia de las propias dificultades*, donde no se encuentran diferencias. La constatación de estas diferencias en la cantidad de mediación, sin embargo, no se ha reflejado de forma clara en las puntuaciones obtenidas en el dispositivo dinámico por cada uno de los grupos.

Tabla88. Número de anotaciones efectuadas en cada uno de los grupos de aplicación y correlación con la puntuación dinámica global

<i>Grupos de aplicación</i> <i>Patrones de mediación</i>	(A) CEIP Andalucía	(B) CEIP A.Machado	(C) CEIP El Ruedo	(D) CEIP El Ruedo	(E) CEIP L.de Rueda	(F) CEIP S. Alonso	(G) IES L. Guzmán	(H) IES L. Guzmán	(I) IES M. Pineda	(J) IES Torreblanca	(K) IES Torreblanca	Total
1(a) Profundización en la respuesta	17	223	31	19	3	34	70	30	11	30	29	497
2 (b) Ampliación de la actividad	6	13	8	4	0	48	0	8	10	0	8	105
3 (c) Apoyo en la comprensión...	3	147	51	37	2	23	39	10	13	73	21	419
4 (d) Recapitulación conjunta...	0	4	0	1	0	4	0	0	0	2	1	12
5 (e) Mediación de la discusión ...	8	2	0	0	1	16	8	9	5	5	2	56
6 (f) Repetición de una actividad...	1	19	9	5	0	4	2	4	0	6	7	57
7 (g) Conciencia sobre la estructura...	0	66	12	0	0	2	4	0	9	0	6	99
8 (h) Conciencia sobre las propias est...	6	60	1	0	0	43	6	18	4	9	8	155
9 (i) Conciencia de lo que sabe...	12	77	10	7	0	18	19	3	8	14	8	176
10 (j) Conciencia de la estructura fon...	2	30	13	7	1	4	0	0	14	0	7	78
11 (k) Conciencia de las propias dif.	0	1	0	0	0	1	0	0	2	0	0	4
12 (l) Mediación de est. de planif.	0	47	1	1	0	22	8	2	1	2	4	88
13 (m) Mediación de est. de superv.	4	58	43	32	6	33	15	20	10	36	9	266
14 (n) Mediación de est. de autoev.	20	78	14	8	7	58	17	16	10	26	14	268
15 (ñ) Mediación de est. de transf..	0	4	0	3	0	26	0	0	0	4	0	37
16 (o) Mediación de est. de int. textual	14	160	25	40	11	27	7	14	13	4	11	326
17 (p) Mediación de est. de int. tex-con.	0	79	0	11	0	18	2	0	0	1	1	112
18 (q) Mediación de la motivación...	2	16	3	2	0	10	0	4	2	27	8	74
19 (r) Mediación de la participación...	7	2	0	0	3	15	1	18	6	5	2	59
<i>Total de anotaciones en cada grupo</i>	102	1086	221	177	34	406	198	156	118	244	146	2888
Valor de correlación	-0,035	-0,369	-1,000**	-1,000**	---	---	---	---	0,148	---	-0,250	
PUNTUACIÓN DINÁMICA GLOBAL	108,936	85,326	75,519	74,082	73,779	105,668	96,583	101,285	95,790	121,887	120,341	

Estructura factorial

A pesar de que la estructura del sistema de categorías ha venido dada en parte, tal y como hemos descrito, por la propia estructura interna del dispositivo de evaluación dinámica, conformándose en torno a los procesos implicados en la lectura que eran contemplados en la propuesta, hemos considerado oportuno someter los datos obtenidos de las anotaciones efectuadas a un *análisis de componentes principales* con el objetivo de analizar la estructura factorial. La técnica empleada permite transformar el conjunto interrelacionado de variables observadas en diversos subconjuntos de variables, denominadas factores, que no mantienen correlaciones entre sí. De entre las diversas opciones que pueden ser tomadas para el análisis factorial, elegimos la solución «re-escalada» del análisis de componentes principales, con rotación varimax. Debido a que la muestra estaba compuesta por diversos grupos con diferentes valores de varianza para cada variable observada, el análisis de las covarianzas fue la medida adoptada para extraer los componentes principales, en detrimento del análisis de las correlaciones. En este caso, los autovalores iniciales resultantes para cada uno de los factores extraídos, deben poseer un valor superior a la varianza promedio de los elementos. Los resultados muestran una estructura compuesta por 11 factores que de forma conjunta explicaban el 79,39% de la varianza.

Mediante un análisis de *conglomerados jerárquicos*, y en función de la estructura de correlaciones existente en la matriz de datos, hemos organizado los factores resultantes. Este procedimiento intenta identificar grupos relativamente homogéneos de variables basándose en características previamente seleccionadas. De esta forma, hemos identificado y agrupado las unidades de información en función de la fuerza de sus correlaciones y, posteriormente, tras obtener los coeficientes de correlación inter-ítems a través del procedimiento para analizar la fiabilidad de los distintos agrupamientos realizados, hemos asignado las unidades de información a los componentes obtenidos. En la tabla 89 resumimos este proceso y presentamos los datos. Comenzamos a continuación exponiendo una descripción de cada uno de los factores:

Factor 1. Hace referencia a conductas de mediación caracterizadas principalmente por la profundización en la respuesta del alumno, realizando nuevas preguntas en función de sus respuestas o su conducta, y estableciendo gradualmente apoyo y/o sugerencias que acerquen al sujeto a la respuesta deseada. Explica el 20,22% de la varianza total. Las unidades de información correlacionadas bajo este factor son: a.pro1; c.est7; c.est11; g.est22; h.con45; i.sab56; l.pla82; m.sup96; n.aut107; n.aut110.

Factor 2. Este factor es con diferencia el que aglutina un mayor número de elementos. Aunque éstos corresponden a diversas categorías de las que hemos considerado previamente, poseen el denominador común de hacer referencia a la mediación del conocimiento metacognitivo y de las estrategias de comprensión, especialmente las de autorregulación. También correlacionan dentro de este conjunto las unidades que destacan la mediación de la discusión metacognitiva, la recapitulación y reconstrucción colectiva de lo trabajado, así como aspectos relativos a la motivación y la participación. Explica el 9,39% de la varianza, relacionándose con los elementos: b.amp3; c.est5; d.rec14; e.dis17; e.dis18; f.rep19; g.est23; g.est29; g.est34; h.con35; h.con36; h.con37; h.con38; h.con40; h.con41; h.con42; h.con43; h.con44; i.sab58; j.fl61; j.fl68; j.fl71; j.fl72; k.dif75; l.pla76; l.pla78; l.pla80; l.pla81; m.sup88; m.sup91; m.sup93; m.sup94; m.sup102; n.aut109; n.aut111; n.aut112; n.aut113; n.aut114; n.aut115; n.aut116;

Tabla 89. Coeficientes de correlación inter-ítems y coeficientes de fiabilidad para cada uno de los factores obtenidos en el análisis.

Factores	Media de anotaciones	Correlaciones inter-ítem	Alfa	Nº de elementos
F1 Profundización en la respuesta y establecimiento de apoyo gradual mediante la realización de preguntas	22,74	0,62	0,63	10
F2 Mediación del conocimiento metacognitivo y de las estrategias de comprensión (autorregulación)	22,46	0,45	0,95	58 ¹
F3 Apoyo, guía y supervisión en la realización de la tarea, con cesión progresiva del control	16,00	0,60	0,90	22
F4 Mediación de la conciencia sobre las propias estrategias de comprensión (estrategias de planificación)	5,14	0,60	0,69	14
F5 Mediación de la toma de conciencia sobre las estrategias necesarias para comprender mejor	4,57	0,46	0,91	20
F6 Mediación de la participación, la discusión, la reflexión y valoración, y el apoyo entre compañeros	4,26	0,67	0,90	8
F7 Mediación de la conciencia estructural y del uso de estrategias relacionadas con la estructura textual	2,49	0,52	0,86	14
F8 Mediación de la conciencia sintáctica y estructural y del uso de estrategias de integración textual	1,66	0,59	0,85	13
F9 Mediación de la conciencia sintáctica (signos de puntuación) y de estrategias de autoevaluación	1,46	0,65	0,69	6
F10 Mediación de la supervisión, integración textual e integración texto-conocimientos (hipótesis)	1,40	0,60	0,90	8
F11 Mediación de estrategias de autoevaluación (supervisión y evaluación de la propia comprensión)	0,34	0,77	0,82	4
<i>Total</i>	<i>82,52</i>	<i>---</i>	<i>0,92</i>	<i>177</i>

¹Aunque el análisis de conglomerados jerárquicos asigna el elemento l.pla80 a este conjunto de variables (factor 2), los cálculos posteriores no lo tienen en cuenta, ya que a la postre no recibió ninguna anotación.

n.aut117; n.aut118; n.aut119; ñ.tra132; ñ.tra133; ñ.tra134; o.tex136; o.tex138; o.tex139; o.tex140; o.tex143; o.tex155; p.con156; p.con157; q.mot171; q.mot172; r.par176; r.par177.

Factor 3. Este factor hace referencia especial a aquellas conductas mediadoras que implican un seguimiento y apoyo cercanos por parte del aplicador, para ir de forma progresiva cediendo el control al alumno. Algunas de estas conductas se concretan en el establecimiento de apoyo en la comprensión y en la realización de la tarea, acompañando al alumno en la ejecución, estableciendo una cesión progresiva del control, favoreciendo una lectura más reflexiva y un mayor control y autorregulación. Se trata en definitiva de comentarios, sugerencias o cuestiones que el aplicador realiza tratando de «acompañar» al alumno en la ejecución de la tarea. Habitualmente, un seguimiento de este tipo, que implica en ocasiones una realización graduada de la tarea junto al sujeto, suele emplearse cuando existen notables dificultades por parte del alumno para realizar de forma autónoma la actividad. Explica el 8,37% de la varianza y se relaciona con los siguientes elementos: c.est4; c.est6; c.est8; c.est9; c.est10; c.est13; f.rep21; h.con46; h.con47; i.sab57; m.sup89; m.sup95; m.sup99; m.sup101; n.aut121; n.aut122; n.aut124; n.aut125; n.aut126; n.aut127; ñ.tra135; q.mot173.

Factor 4. Hace referencia sobre todo a la mediación del conocimiento sobre las propias estrategias de comprensión, especialmente de aquellas que tienen que ver con la

planificación de la lectura. Explica el 7,22% de la varianza y se relaciona con los elementos: g.est31; h.con50; h.con51; h.con53; j.fl62; l.pla77; l.pla84; l.pla86; l.pla87; m.sup105; n.aut130; o.tex145; o.tex149; o.tex154.

Factor 5. Este factor reúne una serie de conductas de mediación que se encuentran repartidas entre los distintos patrones metodológicos contemplados en la matriz. El denominador común de la mayoría de estas conductas radica en favorecer la toma de conciencia sobre qué tipo de estrategias sería necesario aplicar para comprender mejor, ya sea mediante la realización de cuestiones, la explicación bidireccional, la petición al alumno de que realice determinada actividad, o la demostración de un modelo por parte del aplicador. Explica el 6,47% de la varianza y se relaciona con los siguientes elementos: c.est12; g.est30; h.con48; h.con52; h.con54; i.sab59; i.sab60; j.fl67; j.fl70; l.pla79; l.pla83; l.pla85; m.sup106; n.aut131; o.tex146; o.tex148; p.con158; p.con163; p.con164; p.con165.

Factor 6. Hace referencia a las conductas de mediación dirigidas a favorecer la participación activa de los sujetos, la discusión metacognitiva, la reflexión y valoración de su actividad, así como el apoyo entre los compañeros del grupo en la explicación, resolución o valoración de una actividad. Explica el 5,70% de la varianza y se relaciona con los elementos: e.dis16; h.con39; n.aut120; o.tex137; q.mot170; r.par174; r.par175; r.par178.

Factor 7. Los elementos que correlacionan entre sí conformando este factor, hacen referencia básicamente a la mediación de la conciencia sobre los elementos estructurales. Así mismo, también correlacionan otras variables, que aluden a la mediación del uso de estrategias de supervisión (uso del contexto para comprender), y de estrategias de integración textual e integración texto-conocimientos previos, que tienen a su vez, estas últimas, relación con la estructura de los textos. Este factor explica el 5,55% de la varianza total, relacionándose con los siguientes elementos de la matriz: b.amp2; g.est24; g.est25; g.est26; g.est27; g.est28; j.fl69; k.dif74; m.sup104; o.tex153; p.con159; p.con166; p.con167; q.mot169.

Factor 8. Se trata de una serie de elementos, relativamente dispersos, que hacen referencia a aspectos relacionados sobre todo con la conciencia de la estructura sintáctica, el conocimiento sobre las propias estrategias de integración textual, y con la mediación del uso de estas estrategias. Explica el 4,83% de la varianza y se relaciona con los elementos: d.rec15; f.rep20; g.est32; g.est33; h.con49; h.con55; j.fl73; m.sup100; o.tex147; o.tex150; o.tex151; o.tex152; p.con168.

Factor 9. Hace referencia a la conciencia sintáctica, especialmente acerca de los signos de puntuación. También se refiere a la mediación de procesos de supervisión y autoevaluación del alumno acerca de la dificultad de los textos. Explica el 4,65% de la varianza total y se relaciona con los siguientes elementos: j.fl63; j.fl64; j.fl65; m.sup98; n.aut108; n.aut123.

Factor 10. Hace referencia a la mediación de estrategias de supervisión de la propia comprensión, de estrategias de integración textual y de estrategias de integración texto-conocimientos previos, especialmente relacionadas con la formulación de hipótesis sobre el contenido y la estructura del texto. Explica el 3,96% de la varianza total y se

relaciona con los elementos: m.sup92; m.sup97; o.tex141; o.tex142; o.tex144; p.con160; p.con161; p.con162.

Factor 11. Hace referencia a la mediación de algunas estrategias de autoevaluación de la comprensión. Explica el 3,04% de la varianza y se relaciona con los elementos: j.fl66; m.sup103; n.aut128; n.aut129.

Hay que hacer notar que el conjunto de unidades de información que integran cada uno de los factores no es en sí un conjunto totalmente homogéneo. Como podemos comprobar, el grupo de elementos resultante no siempre parece pertenecer a un mismo tipo de procesos o habilidades y en ocasiones no ha sido fácil etiquetar los factores al estar conformados por diversos tipos de variables que, en principio, son indicativos de conductas de mediación que se dirigían a distintos procesos. La estructura de correlaciones de la matriz de datos, en la que se basa el análisis factorial realizado, es compleja y, en no pocas ocasiones, algunos elementos mantienen varias correlaciones positivas y muy significativas. El diseño resultante del análisis factorial es fruto de la elección de unas y la desestimación de otras, en función básicamente de la intensidad de estas correlaciones. Unido a esto, nos parece importante destacar que el sistema de categorías construido y utilizado para el análisis de las anotaciones de los aplicadores, no queda en absoluto invalidado con la extracción de los factores. Estos factores responden, como hemos apuntado, a la estructura interna de correlaciones entre la totalidad de anotaciones efectuadas en función de las unidades de información consideradas. En un estudio sobre patrones metodológicos puestos en práctica en un dispositivo de ED, el hecho de que una determinada conducta mediadora correlacione significativamente con otra, no debería significar necesariamente que ambas conductas pertenezcan a una misma categoría de patrones de mediación; nosotros hemos optado desde el comienzo por clasificar los patrones metodológicos en función de los procesos implicados en la lectura que eran considerados en nuestra propuesta de evaluación dinámica, y que, a su vez, respondían a la teoría de base adoptada. Estas categorías han ido conformándose y enriqueciéndose con las anotaciones de los aplicadores. En este sentido, no es ilógico; antes bien, esperable, que diversas conductas de mediación, enfocadas en un principio a distintos procesos de lectura, puedan ponerse en práctica de forma inter-relacionada. La estructura factorial resultante del presente análisis coincide notablemente con los patrones metodológicos establecidos, pero más allá de esto, su esencia misma agrupa aquellas conductas mediadoras que mantienen correlaciones entre sí y que, a su vez, no las mantienen —o al menos no con la misma intensidad— con otras conductas de otros conjuntos. Por otra parte, los resultados del análisis nos ofrecen información relevante y nueva sobre las categorías establecidas y sobre las relaciones que las unidades de información contenidas en dichas categorías mantienen entre sí y con otras unidades de otras categorías.

Análisis y valoración conjunta del Grupo de Trabajo y del dispositivo de evaluación dinámica EDPL

En el capítulo dedicado a exponer el procedimiento seguido en nuestra investigación, hacíamos referencia al proceso y la temporalización llevados a cabo en las sesiones del Grupo de Trabajo (GT). Hacíamos alusión entonces a la relevancia del método seguido en la formación de los profesores aplicadores, método caracterizado por la discusión y el análisis conjunto, que permitía conocer con mayor profundidad el contenido de la propuesta metodológica de evaluación, al tiempo que unificar criterios en el uso y

puesta en práctica de los patrones metodológicos de mediación y en la evaluación a través de los indicadores propuestos. Decíamos también que al finalizar el curso se realizó una valoración del proceso de trabajo seguido en el GT, así como del aprovechamiento del mismo durante el periodo de funcionamiento. Esta valoración se realizó a través de un documento aportado por el CEP que debía reflejar una *Memoria* del trabajo desarrollado durante las sesiones de trabajo. De este documento para la valoración aportado por el CEP, y cumplimentado mediante el análisis y la valoración conjunta de los miembros del GT, hemos extraído aquellos puntos más relevantes y que nos sirven para acometer la valoración cualitativa de los aplicadores sobre el trabajo desarrollado. Para secuenciar este análisis, utilizamos los epígrafes del documento base de valoración.

Análisis y valoración conjunta de la labor desarrollada en el Grupo de Trabajo

Grado de satisfacción de los integrantes del grupo con relación al trabajo desarrollado. Objetivos logrados y su incidencia en la práctica de aula o del Centro. Objetivos que aún no se han conseguido.

Los objetivos del GT contemplaban el conocimiento, la aplicación y la evaluación de un dispositivo de evaluación dinámica de la comprensión de textos en alumnado que presenta dificultades en el aprendizaje de la lectura. Pretendíamos, al tiempo que proporcionar un espacio reflexión y discusión sobre la forma más idónea de abordar la evaluación y la intervención para la mejora de la lectura y la comprensión, comprobar si la aplicación del citado dispositivo mejoraría el rendimiento del alumnado al cual se aplicaría.

Lógicamente, al finalizar las sesiones del GT, aun no podíamos analizar en profundidad la consecución de todos los objetivos propuestos; pero si habíamos recogido suficiente información para ofrecer algunos datos de carácter cualitativo. Esta información se relacionaba sobre todo con la valoración del dispositivo trabajado por parte de los profesores integrantes del GT, aunque también se hacía referencia a observaciones sobre el aprendizaje del alumnado, sobre las estrategias de enseñanza desarrolladas o sobre las condiciones de aplicación del dispositivo.

El gradiente que expresaría el nivel de satisfacción de los integrantes del GT es amplio. En general, la mayoría de las ideas y de los elementos destacados durante las sesiones dedicadas a la valoración del trabajo desarrollado, expresaban un grado de satisfacción muy positivo, aunque también se hizo referencia a diversas dificultades, especialmente relacionadas con las circunstancias en las que se desarrolló la aplicación práctica. Las ideas destacadas durante las sesiones de valoración fueron las siguientes:

- Satisfacción por el trabajo desarrollado y por la observación en el alumnado de mejoras en los procesos trabajados:
 - *"Estoy muy satisfecha por el trabajo realizado... En los niños estoy viendo el fruto del trabajo."*
 - *"A la hora de afrontar las actividades, no lo hacen igual que antes... En general lo hacen mejor."*
 - *"Veó bien el trabajo, y a los niños les ha servido bastante."*

- "Van contentos a las actividades."
- "He observado mejoras en los procesos de atención; los niños ahora preguntan más con relación a la tarea que deben realizar y se muestran mas concentrados cuando comienza el trabajo."
- Constatación de haber observado como profesores/as, aspectos relativos a la evaluación de los procesos de comprensión que antes no eran observados.
- Alusión al material aportado, en el sentido de que su conocimiento y aplicación ha supuesto una oportunidad de aprendizaje. Se observa el manual y la propuesta metodológica recogida en este, como un recurso valido para:
 - Desarrollar estrategias de autorregulación y sugerir al alumnado la puesta en práctica de estas estrategias.
 - Observar niveles de apoyo que cada niño/a necesita.
- Constatación de que el trabajo desarrollado ha aportado estrategias nuevas, relacionadas con el afrontamiento de la evaluación y la intervención en las dificultades de aprendizaje de la lectura.
- Alusión al mayor aprovechamiento que ahora se extrae de la realización de una actividad de lectura.
- Constatación de haber ganado experiencia durante este curso a través del trabajo desarrollado con la puesta en practica del dispositivo de evaluación dinámica: "*Estoy segura de que el curso que viene lo haré mejor...*"
- Alusiones al aprendizaje y aprovechamiento obtenidos a través de las sesiones de análisis conjunto de los videos sobre la aplicación del dispositivo, así como de las sesiones en las que los/as compañeros/as han expuesto sus anotaciones sobre la aplicación de las actividades.
- Alusión al aprendizaje que ha supuesto el trabajo desarrollado como *guía* en la observación de las dificultades de comprensión en el alumnado; se destaca la utilidad de contar con unos indicadores precisos en los que fijarse a la hora de evaluar e intervenir sobre la lectura.
- Alusión a las dificultades que ha supuesto la realización de algunas de las actividades propuestas.
- Alusión a las dificultades provenientes de la necesidad de aplicación colectiva del dispositivo, debido a las condiciones en las que tuvo lugar el trabajo de apoyo de los profesores integrantes del GT. Estas condiciones han dificultado principalmente el registro de la información considerada relevante en la evaluación de los procesos implicados, así como una idónea valoración del *proceso* de resolución de las actividades de lectura para cada uno de los alumnos.
- Alusiones a las dificultades provenientes del horario, ya que, en ocasiones, se ha debido compaginar la aplicación del dispositivo con otras labores de apoyo a otros alumnos presentes también en el aula.
- Se alude también por parte de algunos integrantes del GT al problema que ha supuesto la discontinuidad en la aplicación, ya que, en ocasiones, se han debido interrumpir las actividades por la necesidad de realizar otras ya programadas por el Centro.
- Constatación de las dificultades provenientes del trabajo con algunos de los textos propuestos en el manual.

- Insatisfacción por las condiciones en las que se ha aplicado, relacionadas con la aplicación colectiva: *"No estoy demasiado satisfecha... Un niño ha hecho que el trabajo se ralentice demasiado, no pudiendo avanzar con todo el grupo. El resto del grupo tenía otro nivel de comprensión y quizás no han aprovechado bien las actividades que hemos trabajado. Han realizado poco trabajo y no hemos llegado a realizar las actividades que mas hubieran interesado."*
- Alusión a las dificultades provenientes de la desmotivación en algunos grupos de alumnos/as, así como de la falta de continuidad en el trabajo debido a aspectos como el absentismo y las expulsiones de alumnado participante en E.S.O.
- Alusión a aspectos relacionados con las dificultades de coordinación en los centros. Estas dificultades se destacan con relación a la realización de propuestas para generalizar el trabajo desarrollado. Existe, según la mayoría del profesorado participante, la dificultad de que la forma de trabajar propuesta y desarrollada durante el curso, sea asumida por el centro en su conjunto.
- Constatación de que el trabajo desarrollado, en la mayoría de los casos, no ha trascendido del aula de apoyo. *"Nadie ha preguntado nada..."*

En este sentido es necesario decir que no era un objetivo del GT el generalizar la propuesta metodológica trabajada, relativa a la evaluación-intervención en las dificultades de comprensión, al resto del profesorado de los centros participantes; este objetivo superaba con creces nuestras posibilidades; sin embargo, durante el transcurso de las sesiones de evaluación del GT, sí ha sido éste un tema que preocupaba a los integrantes del grupo, en la medida en que la propuesta se observaba como idónea para incrementar la eficacia del trabajo de apoyo que se realiza en los centros con relación a las dificultades de aprendizaje de la lectura.

En línea con lo anterior, se realizaron propuestas por parte de algunos integrantes del GT, como aplicar algunas de las actividades en el aula ordinaria para hacer participe al tutor del contenido y la metodología de las mismas. En este sentido se hablaba de: *"hacer participe al tutor de la evaluación psico-pedagógica encaminada a la valoración de las dificultades de comprensión en el alumnado."* También se proponía la formación de otros Grupos de Trabajo para generalizar la labor desarrollada durante la aplicación experimental del dispositivo.

- Constatación de las limitaciones de la influencia y del trabajo desarrollados en el alumnado, debido a la falta de coordinación y coherencia en el trabajo educativo dirigido a la mejora de la lectura en los centros.

Descripción y puesta en práctica del trabajo desarrollado. Valoración del proceso y de los resultados

El GT comenzó a trabajar con un material que ya estaba elaborado. Como ya hemos indicado, se trataba de la aplicación experimental de un dispositivo de evaluación dinámica de procesos implicados en la lectura. Este dispositivo, así como los dos documentos de formación entregados al profesorado participante o las *Hojas de registro y evaluación* para las anotaciones sobre la aplicación de las actividades de evaluación con cada sujeto, han constituido la base sobre la cual se desarrolló la labor del GT. Dado que la descripción y el análisis del dispositivo en su conjunto, ha sido ya realizada, nos

disponemos ahora a valorar la puesta en práctica del trabajo desarrollado por el grupo durante las sesiones de trabajo.

La puesta en práctica de este trabajo por parte de los integrantes del GT ha supuesto el análisis y la contextualización de las propuestas contenidas en el mismo. Concretamente, la producción material del GT durante el curso ha consistido en:

- Complimentación de las *Hojas de registro y evaluación* para cada uno de los alumnos o alumnas participantes. En estas, el profesor debía analizar el proceso de resolución de las actividades de evaluación realizadas, exponer el proceso de mediación efectuada, para obtener la máxima información posible sobre las dificultades de aprendizaje y sobre las posibles orientaciones para paliar esas dificultades. Por último debía analizar la puesta en práctica de procesos metacognitivos por parte del alumno durante la realización de las actividades.

En cuanto al registro efectuado por los integrantes del GT, ya hemos hecho mención de que las distintas condiciones de aplicación del dispositivo de evaluación dinámica, han conllevado igualmente diferencias en cuanto al modo de registrar la información sobre la actividad desplegada por el alumnado. En ocasiones se han llevado *registros colectivos e individuales* en función de las posibilidades que se tenían de trabajar con un grupo numeroso o con un alumno en particular. Los registros colectivos han servido posteriormente de base para volcar en cada una de las *Hojas de registro y evaluación* la información relevante sobre el proceso de mediación y evaluación realizado, añadiendo para cada alumno la información específica que el profesor consideraba oportuna, en el caso de que pudiera haber sido recogida.

- Anotaciones a modo de diario sobre cada una de las sesiones de trabajo con el alumnado, anotaciones que servían posteriormente para complimentar las hojas de registro.
- Grabaciones en video de algunas sesiones de trabajo, para favorecer una validación de las observaciones por parte de distintos observadores y también con objeto de contribuir a la formación de los integrantes del GT, mediante el análisis conjunto de los patrones de mediación y evaluación de procesos metacognitivos en las sesiones de evaluación.

En cuanto a la valoración global del proceso y de los resultados obtenidos con la puesta en práctica del dispositivo de evaluación dinámica, esta se considera especialmente positiva en los aspectos relativos al propio aprendizaje del profesorado participante en el GT (aprendizaje de estrategias de evaluación e intervención y aprendizaje teórico sobre el proceso de comprensión y los procesos implicados), así como a la intuición-constatación de los beneficios que parece haber tenido su aplicación para el alumnado participante.

Expectativas de continuidad del GT. Línea de trabajo a seguir en el futuro.

Con relación a las expectativas de continuidad del GT, es destacable que la mayoría de los participantes expresaron su deseo e intención de continuar la puesta en práctica de

las actividades del dispositivo el próximo curso. Esta continuidad en el trabajo planteado con algunos alumnos y alumnas era considerada especialmente idónea por algunos profesores, aunque existía la dificultad práctica de la posible no continuidad de los mismos en las plazas que ocupaban. Por otra parte, hemos comentado que, a efectos del estudio planteado al inicio de un curso escolar —estudio que servía de base a los trabajos desarrollados en el GT—, la mayor parte de la información que se pretendía recoger, se habría obtenido ya al finalizar ese mismo curso. Sin embargo, se ha planteado, sin olvidar las dificultades expuestas anteriormente, que sería idónea la continuidad del GT para completar la aplicación del dispositivo y seguir aprendiendo del análisis conjunto realizado por sus miembros. En cualquier caso, factible o no dar continuidad al trabajo emprendido mediante este formato de GT, la intención de sus integrantes es seguir manteniendo un contacto encaminado fundamentalmente a completar información cuantitativa y cualitativa sobre el proceso de evaluación desarrollado, así como a compartir y valorar los resultados que el análisis de datos que ha seguido a la recogida de información, haya producido.

Como sugerencias prácticas para una posible continuidad del trabajo desarrollado en el grupo, se expusieron las siguientes:

- Centrarnos en la aplicación de una sola actividad. Todos los integrantes del GT aplicarían al mismo tiempo una actividad del dispositivo y se analizaría conjuntamente en las sesiones. Esto ayudaría a profundizar en los distintos elementos metodológicos y de fundamentación teórica recogidos en el dispositivo.
- Analizar grabaciones. Aunque esto se ha podido realizar en parte, la constatación por parte del grupo de su utilidad para valorar los procesos de mediación y también como elemento de formación para el propio grupo, ha hecho que se valore muy positivamente este elemento entre los integrantes del grupo.
- Centrarnos especialmente en el *proceso* de aplicación del dispositivo. Esto es algo sobre lo que ya durante las sesiones del GT se ha venido trabajando. Tanto las anotaciones de los profesores como el contenido del análisis conjunto que se realizaba en las sesiones de trabajo, eran continuamente conducidas y reconducidas hacia el proceso mismo de mediación y de evaluación de procesos metacognitivos en la resolución de las actividades de lectura.

Por otra parte, también en el capítulo de *procedimiento* se hace referencia a la valoración conjunta de carácter cualitativo que efectuaron los integrantes del GT sobre la puesta en práctica del dispositivo EDPL, de forma especial sobre: (a) los patrones metodológicos de mediación y evaluación dinámica para valorar los procesos metacognitivos implicados en la lectura, (b) la utilidad y eficacia del procedimiento dinámico seguido para recabar información valiosa para la mejora del proceso de comprensión, así como sobre (c) las dificultades encontradas en su aplicación. Esta valoración conjunta se realizó a raíz de la cumplimentación individual de un pequeño cuestionario de preguntas abiertas, que posteriormente sería analizado en el grupo. Hemos procedido al análisis cualitativo de las respuestas otorgadas por los integrantes del GT, análisis que exponemos a continuación siguiendo las cuestiones formuladas. En la parte izquierda de la tabla hemos anotado aquellas unidades de información destacadas por los aplicadores en sus respuestas, mientras que en la parte derecha hemos

recogido el número de veces que fueron aludidas dichas unidades. Hemos categorizado, así mismo, las unidades de información con el fin de obtener una percepción más clara y útil de los elementos destacados. El orden de exposición para cada una de las cuestiones formuladas es decreciente, de tal forma que figuran antes aquellos aspectos que fueron destacados en mayor número por los aplicadores.

Valoración del dispositivo de evaluación dinámica de procesos implicados en la lectura (EDPL) y de su aplicación

Cuestión 1

Principales dificultades encontradas en la aplicación del dispositivo	<i>Número total de anotaciones que aludían a estas dificultades</i>
<p>Dificultades relacionadas con aspectos organizativos del aula o del centro o con las circunstancias socio-culturales del entorno</p> <p>Falta de continuidad/sistematización en la aplicación debido a:</p> <ul style="list-style-type: none"> • horarios del alumnado de apoyo • organización de los grupos de apoyo y aspectos organizativos del centro en general • necesidad de retomar la actividad después de varios días • ausencias del alumnado (absentismo) y expulsiones • necesidad de atender a otros alumnos no participantes del estudio 	13
<p>Dificultades derivadas de la aplicación de un procedimiento dinámico de evaluación</p> <ul style="list-style-type: none"> • registrar la información requerida durante el proceso de aplicación • la observación/evaluación de algunos indicadores • falta de conocimientos sobre la aplicación del dispositivo 	5
<p>Dificultades relacionadas con la aplicación grupal del dispositivo</p> <ul style="list-style-type: none"> • dificultad en la mediación y en el registro de la información • necesidad de aplicación individual de algunas actividades, lo que implicó un enlentecimiento de todo el proceso y situaciones poco «dinámicas» 	4
<p>Dificultades relacionadas con el contenido de EDPL</p> <ul style="list-style-type: none"> • falta de adecuación de algunos textos (nivel alto) • falta de organización del material aportado 	3
<p>Dificultades relacionadas con el alumnado</p> <ul style="list-style-type: none"> • falta de motivación del alumnado 	1
<p>Dificultades relacionadas con el contexto escolar</p> <ul style="list-style-type: none"> • divergencias entre la propuesta dinámica realizada y el trabajo habitual desarrollado en los centros con relación a la lectura • la aplicación no ha tenido incidencia en el trabajo desarrollado en el aula ordinaria 	1

Cuestión 2

Elementos del dispositivo que podrían mejorarse	<i>Número total de anotaciones que aludían a estos elementos</i>
<p>Mejoras relacionadas con el contenido de las actividades</p> <p>Presentación de los textos y las actividades</p> <ul style="list-style-type: none"> • presentación más atractiva de los textos y las actividades (dibujos, distintas tipologías de letras) <p>Adecuación del contenido</p> <ul style="list-style-type: none"> • introducir textos con nivel más bajo • elaborar una propuesta de actividades para primaria y otra para secundaria • distintos niveles de complejidad en las tareas o mayor cantidad <p>Dificultad y amplitud de las tareas</p> <ul style="list-style-type: none"> • bloque de actividades de asociación G-F excesivamente largo y dificultoso (actividades 11 y 12) 	11
<p>Mejoras relacionadas con el proceso de evaluación</p> <ul style="list-style-type: none"> • recoger los indicadores de evaluación en el Cuaderno de Hojas de registro y evaluación • sistematizar más el contenido de la propuesta de evaluación de los procesos de autorregulación (normalización del proceso de mediación) • contar con más espacio en el Cuaderno de Hojas de registro y evaluación para realizar las anotaciones • en aplicaciones grupales, poder recoger la información de forma colectiva, teniendo la posibilidad de destacar aspectos concretos de algún alumno • contar con versiones similares de una misma actividad para poder evaluar la evolución en la práctica de las estrategias 	6
<p>Mejoras en la teoría de base del dispositivo</p> <ul style="list-style-type: none"> • explicitar más los microprocesos implicados en cada una de las tareas 	1

Cuestión 3

Elementos del dispositivo considerados <i>clave</i> para obtener información valiosa para una posterior intervención	<i>Número total de anotaciones que aludían a estos elementos</i>
<p>Elementos relacionados con la propuesta metodológica</p> <ul style="list-style-type: none"> • la evaluación dinámica y la mediación • la propuesta permite seguir un método de trabajo y sacar más provecho a la realización de las actividades • centrarse en los indicadores de evaluación • la puesta en práctica de los patrones de mediación 	11

Elementos relacionados con el registro de la información	
<ul style="list-style-type: none"> • el registro de la información en el Cuaderno de Hojas de registro y evaluación • las anotaciones realizadas a modo de diario • grabaciones en video de algunas sesiones 	5
Elementos relacionados con los procesos contemplados en EDPL	
<ul style="list-style-type: none"> • los procesos implicados en cada una de las tareas • los procesos de autorregulación (microprocesos y macroprocesos) • la explicación de los procesos implicados 	3
Elementos relacionados con los aplicadores	
<ul style="list-style-type: none"> • lectura y reflexión sobre las actividades • puesta en común/análisis colectivo de la puesta en práctica 	2
Elementos relacionados con la propuesta de actividades	
<ul style="list-style-type: none"> • la realización de las actividades ofrece información que orienta la posterior intervención 	1

Cuestión 4

Elementos del proceso seguido de <i>Evaluación dinámica-Intervención</i> considerados <i>clave</i> para explicar la posible mejora en los procesos de lectura	<i>Número total de anotaciones que aludían a estos elementos</i>
Elementos relacionados con los procesos trabajados	
<ul style="list-style-type: none"> • las estrategias trabajadas con el alumnado • el trabajo sistemático sobre las habilidades metacognitivas • la intervención sobre los distintos procesos: metacognición, integración textual, etc. • favorecer la toma de conciencia sobre las estrategias trabajadas para realizar con éxito las tareas 	7
Elementos relacionados con la propuesta metodológica	
<ul style="list-style-type: none"> • la mediación ejercida por los aplicadores • la propuesta metodológica, centrada en la evaluación-intervención en procesos metacognitivos • la mediación de procesos metacognitivos • detenerse (lentificar el proceso) en los bloques de actividades donde se hallaron dificultades • realizar la aplicación grupal de forma lúdica y participativa, siguiendo algunas de las pautas de interacción del programa CyT 	7
Elementos relacionados con el esquema de actividades propuesto	
<ul style="list-style-type: none"> • la secuenciación y programación de las actividades propuestas • la realización de cada una de las actividades propuestas siguiendo las pautas metodológicas • el seguimiento de la galaxia de actividades 	4

Cuestión 5

Circunstancias que han influido o influirían positivamente en una mayor eficacia o utilidad de la aplicación del dispositivo	<i>Número total de anotaciones que aludían a estas circunstancias</i>
Circunstancias relativas al trabajo de formación de los aplicadores <ul style="list-style-type: none"> • análisis colectivo de las grabaciones en video de las sesiones • las aportaciones de los compañeros sobre la puesta en práctica del dispositivo • mejor y mayor conocimiento de la aplicación del dispositivo y de sus actividades (uso de grabaciones en video) 	6
Circunstancias relativas al contexto escolar <ul style="list-style-type: none"> • generalizar la aplicación al aula ordinaria • coordinación de objetivos y pautas de intervención con el resto de compañeros que atiende al alumno o al grupo evaluado • la colaboración de los tutores • perspectiva metacognitiva en el trabajo desde cualquier área 	5
Circunstancias relativas al tiempo empleado o al momento de la aplicación <ul style="list-style-type: none"> • no estar condicionados por el tiempo (evaluación post-tratamiento) • haber podido repetir algunas actividades del mismo tipo a las realizadas • el trabajo simultáneo de evaluación-intervención durante al menos dos cursos escolares • aplicación sistemática • la hora de aplicación (mejor en la primera hora) 	5
Circunstancias relativas al número de alumnos <ul style="list-style-type: none"> • aplicación individual de algunas actividades • aplicación individual, sin tener que atender otros alumnos • reducir el número de alumnos 	4
Circunstancias relativas al contenido del dispositivo <ul style="list-style-type: none"> • el componente lúdico y divertido de las actividades (los alumnos se han sentido más motivados) 	1

Enriquecimiento de la propuesta mediante la ampliación de las actividades

Como hemos ya comentado, se procuró que la aplicación experimental del dispositivo EDPL fuera realizada siguiendo una cierta línea de actuación que igualara en lo posible las condiciones de su puesta en práctica. Para lograr esto, se incidió de forma especial durante las primeras sesiones de formación en el GT, en la importancia de seguir las pautas metodológicas recogidas en el Manual entregado. Por otra parte, también se pidió a los aplicadores que siguieran un determinado *camino*, trazado igualmente en el Manual, y a través del cual se pretendía que todos los grupos de aplicación realizaran las mismas actividades en el mismo orden. El objetivo de estos intentos de igualar las condiciones de la puesta en práctica, no era otro que incrementar el número de sujetos con condiciones similares a los que poder aplicar análisis estadísticos que fortalecieran las conclusiones del estudio. Sin embargo, también hemos destacado que el dispositivo dinámico desarrollado, tenía originalmente entre sus principales características definitorias la contextualización de los instrumentos de evaluación y la flexibilidad con

la que los evaluadores podían utilizar las actividades de evaluación propuestas. Estas características han seguido presentes en la aplicación experimental, aunque necesariamente su utilización por parte de los aplicadores, se ha visto limitada. Una de las formas en que se ha puesto de manifiesto la contextualización de los instrumentos de evaluación, ha sido la *ampliación de las actividades*. En efecto, durante el transcurso de la aplicación del dispositivo EDPL, la mayoría de los aplicadores ha desarrollado nuevas tareas que complementaban o ampliaban las actividades propuestas en el Manual. Muchas de estas tareas tenían un componente lúdico y, en muchas ocasiones, han sido generadas como fruto de la propia implicación de los alumnos en la actividad desarrollada. No pocas han sido las veces que los propios alumnos han propuesto tareas que se relacionaban con los aspectos que se estaban trabajando. En el volumen de anexos (Anexo XVII) listamos estas tareas que no formaban parte inicialmente de la propuesta metodológica, pero cuyo desarrollo sí responde por entero a nuestra propuesta de evaluación dinámica de los procesos implicados en la lectura.

Algunos ejemplos de la mediación efectuada por los aplicadores

Uno de los aspectos fundamentales del estudio realizado, así como de nuestra propuesta de evaluación dinámica de procesos implicados en la lectura, consiste en el análisis de la interacción mediada del aplicador, así como de la respuesta de los sujetos a la mediación. En el transcurso de este mismo capítulo ya hemos desarrollado el análisis del contenido de las *Hojas de registro y evaluación*, donde los aplicadores han plasmado por escrito el proceso de mediación efectuado, así como sus valoraciones sobre este mismo proceso y sobre la puesta en práctica de los distintos procesos de comprensión implicados. Hemos estructurado el contenido de las anotaciones realizadas por los aplicadores, sirviéndonos de un sistema de categorías que ha ido construyéndose de forma progresiva a raíz de los patrones de mediación recogidos en la propuesta inicial y con las aportaciones de los propios aplicadores en la fase experimental del estudio. Finalmente, hemos extraído de este análisis una serie de *patrones metodológicos de evaluación dinámica de procesos implicados en la lectura*. Ahora queríamos detenernos en algunas de las sesiones de aplicación y mostrar de forma completa, a modo de ejemplo, la transcripción de la interacción mediada y de la respuesta de los sujetos participantes (Anexo XVIII). Se trata de la transcripción fiel de las anotaciones efectuadas por los propios aplicadores en el transcurso de las sesiones de trabajo, por lo que en ocasiones el lenguaje utilizado puede caracterizarse por ser en demasía sintético. Igualmente, debemos señalar que las anotaciones realizadas por los aplicadores han tratado de recoger fielmente las respuestas de los sujetos, por lo que en ocasiones pueden existir errores de expresión. Hemos respetado en todo momento estas anotaciones, situando entre paréntesis algunos comentarios que permiten aclarar o situar algunas de las respuestas. Por otra parte, hemos preferido transcribir el trabajo realizado sobre actividades completas, aunque para ello hayamos debido recoger más de una sesión. De esta forma, pensamos que se refleja mejor la labor de mediación ejercida por el aplicador, al tiempo que puede llegar a observarse la evolución en la propia actividad y en la respuesta del alumnado participante. Las transcripciones aparecen acompañadas de nuestro análisis, efectuado en función de la puesta en práctica de los patrones metodológicos de evaluación dinámica utilizados en el transcurso de la mediación. Hemos tratado de señalar de forma paralela a la transcripción efectuada por el aplicador, aquellas conductas de mediación puestas en práctica, precediéndolas de un asterisco y realizando en ocasiones algunos comentarios relativos al uso o ausencia de algunos de estos patrones metodológicos en la mediación. El contenido íntegro de las actividades a las que se hace referencia en este análisis, tal y como aparecía en el Manual que fue entregado a los aplicadores, puede encontrarse en el volumen de anexos (Anexo XIII).

9.2 Análisis de los cuestionarios EIDAL y EIDAL-MP sobre expectativas e ideas acerca de las dificultades de aprendizaje de la lectura

Aunque un análisis más exhaustivo de la información proporcionada por estos cuestionarios será ofrecido en otra ocasión, queremos sin embargo abordar aún en modo exploratorio algunos de los aspectos derivados del análisis que más tienen que ver con la evaluación dinámica propuesta y realizada. Además de ofrecer algunos datos que pueden resultar de interés para contextualizar los resultados del estudio, es para nosotros igualmente relevante establecer algunas relaciones entre los aspectos recogidos en estos cuestionarios y aquellos relacionados con los procesos y contextos de aprendizaje recogidos a lo largo de la aplicación experimental del dispositivo EDPL.

El procedimiento que hemos seguido para la elaboración de los sistemas de categorías construidos, ha sido similar al expuesto con relación a los Patrones metodológicos de evaluación dinámica utilizados en la mediación. De esta forma, del análisis de las respuestas plasmadas en los cuestionarios, se han extraído una serie de unidades de información a las que han hecho referencia los profesores y las familias, y que posteriormente han sido agrupadas en categorías más inclusivas. Estas categorías son las que aparecen destacadas en negrita, dentro de las tablas que mostramos en el anexo XI-A. Hemos creído oportuno mostrar así mismo las unidades de información, que resumen a su vez grupos de respuestas con idéntico significado. De esta forma, queremos ofrecer la posibilidad de tener un conocimiento más amplio sobre el contenido de las respuestas obtenidas, al tiempo que sobre la propia estructura del sistema de categorías elaborado. Esto permite al interesado compartir en cierto modo el conocimiento del investigador, pudiendo fijar su atención y su análisis en aspectos que no tienen por qué ser los que el autor del estudio ha considerado relevantes.

Por último, antes de abordar el análisis de las respuestas obtenidas en cada uno de los cuestionarios, creemos necesario destacar la posibilidad de que el hecho de haber cumplimentado y entregado los cuestionarios, constituya, ya de entrada, un primer sesgo que, pensamos, debe ser tenido en cuenta. En efecto, es posible que, en su mayoría, las familias y los docentes que han dedicado un tiempo considerable a cumplimentar los cuestionarios, teniendo en cuenta sobre todo que se trataba de responder a cuestiones abiertas que requerían de una cierta reflexión, tengan un grado mayor de implicación o que, de alguna forma, consideren positiva o satisfactoria su labor educativa con relación a los sujetos que presentan dificultades de aprendizaje de la lectura.

Análisis exploratorio del cuestionario EIDAL

Con relación al cuestionario EIDAL (Anexo XI), la muestra finalmente utilizada para la construcción del sistema de categorías que exponemos en el volumen de anexos, ha estado compuesta por 27 docentes. Hemos seleccionado para este análisis, de entre los 31 que cumplimentaron el cuestionario, a aquellos que participaron como profesores o profesoras del alumnado incluido en los grupos experimental (N = 15) o control (N = 12) del estudio. Por otra parte, hemos organizado el sistema de categorías con relación a las siete dimensiones contempladas en el cuestionario en función del contenido de las cuestiones que plantea.

La exposición de este sistema de categorías elaborado, junto con el número de referencias realizadas por el profesorado que participó cumplimentando el cuestionario, puede ofrecernos una información muy valiosa para contextualizar el proceso de aplicación del dispositivo, otorgándonos la posibilidad de «visualizar» de algún modo cómo ha podido tener lugar el trabajo realizado en algunos de los centros en los que se ha aplicado EDPL; así mismo, también puede ofrecernos una información valiosa y adicional sobre algunos de los resultados obtenidos en el estudio. A continuación destacamos algunos puntos que consideramos relevantes con relación a la información obtenida. Lo haremos efectuando un recorrido por los distintos bloques que componen el sistema de categorías propuesto.

En primer lugar, con relación al *origen de las dificultades de aprendizaje y a los factores que influyen en éstas*, podemos observar que el profesorado participante concentra el origen de las dificultades de forma mayoritaria en el propio alumnado. La falta de motivación y de esfuerzo serían aspectos cruciales en este sentido. No obstante, también se otorga importancia al contexto social, familiar y escolar, incluyendo en este último, aspectos relacionados con el proceso de enseñanza-aprendizaje. En línea con las respuestas y argumentaciones encontradas con relación al origen de las dificultades, la ausencia de motivación hacia a la lectura, también se debería en gran medida a factores relativos al propio alumnado; mientras que el contexto escolar figura como el factor menos determinante en la motivación del alumnado, apareciendo, dentro de éste, las alusiones a la presentación atractiva e interesante de los textos y de la lectura, como el factor más relevante.

En cuanto a las dificultades concretas que manifiesta el alumnado con el que trabaja el conjunto de docentes que cumplimentó el cuestionario, las opiniones destacan esencialmente la importancia de los aspectos relacionados con los procesos psicológicos subyacentes, los procesos de asociación G-F y los procesos de integración textual; en cambio, la alusión a los procesos metacognitivos y a los procesos de ajuste P-S, es muy minoritaria, mientras que las referencias a procesos relacionados con la integración texto-conocimientos previos, son inexistentes. Por otra parte, los aspectos que figurarían en la base de estas dificultades, estarían nuevamente relacionados con el propio alumnado; aunque en este caso son importantes también las referencias a los factores relacionados con el contexto escolar (organización escolar, procesos de enseñanza-aprendizaje, y formación del profesorado).

En cuanto a los *aspectos metodológicos relacionados con los procesos de enseñanza/facilitación de la lectura, así como con los procesos de evaluación*, podemos observar en primer lugar un mayor número de referencias por parte del profesorado, que expresan *acuerdo* en tratar, discutir y confrontar ideas previas, contenidos y estrategias de lectura, con relación a las que expresan objeciones o desacuerdo. No obstante, un análisis del contenido de las respuestas otorgadas parece mostrar ciertas limitaciones en la práctica real de estas ideas o en la comprensión total de las implicaciones que conlleva la metodología propuesta en la cuestión. Por otra parte, ante la pregunta relativa a cómo se trabajarían en el aula aquellos aspectos considerados más relevantes en la comprensión de los textos, la enseñanza de estrategias y el análisis/diálogo comprensivo sobre el desarrollo del texto, se sitúan como las acciones más aludidas en la metodología a utilizar.

Por su parte, los elementos destacados como más relevantes a la hora de la enseñanza de la lectura tienen que ver con el proceso de enseñanza y aprendizaje y, sobre todo, con estrategias metodológicas de carácter general, es decir, no centradas en el trabajo sobre procesos concretos implicados en la lectura. Sin embargo, el análisis de las respuestas relacionadas con las actuaciones concretas que se llevaron a cabo para paliar las dificultades de aprendizaje de la lectura, muestra la primacía de aquellas acciones que trataron de incidir sobre el aula, en contraposición a las que se dirigían especialmente a los aspectos relacionados con el centro en su conjunto, y, dentro de las medidas que incidían sobre el aula, destacan las que se dirigían al trabajo sobre un proceso concreto implicado en la lectura. En este caso, las medidas relacionadas con los procesos de ajuste personal-social y aquellas que tenían que ver con los procesos de integración textual, fueron las más utilizadas. Las medidas que trataron de incidir sobre los procesos metacognitivos fueron nuevamente muy minoritarias (tan sólo una referencia).

Con relación a la necesidad de diferenciación o adaptación de la metodología en el proceso de enseñanza de la lectura, en función de la diversidad del alumnado y con el objetivo de lograr avances en su desarrollo, la mayor parte del profesorado piensa que es necesario utilizar métodos adecuados a las necesidades del alumnado. No obstante, también se hicieron referencias que expresaron diversos cuestionamientos a la idoneidad de adaptar los métodos de enseñanza, o que hicieron alusión a las dificultades prácticas para llevar a cabo las adaptaciones. Igualmente, aunque de forma muy minoritaria, se indicó la necesidad de formación del profesorado en este sentido. Por otra parte, la mayor parte de las respuestas dadas con respecto a cómo tiene lugar la evaluación de la comprensión de textos, hacen referencia a la utilización de actividades que, de una u otra forma, implican la lectura de un texto y la realización posterior de actividades (preguntas sobre el texto, realización de resúmenes, análisis y diálogo sobre los personajes, etc.). Sólo encontramos dos referencias explícitas a la evaluación continua, aunque algunas de las propuestas recogidas parecían conllevar también un seguimiento de la actividad del alumnado. Por último, entre las respuestas que se centran en la evaluación de aspectos concretos de la lectura, no encontramos alusiones a la evaluación de procesos metacognitivos.

La tercera dimensión contemplada en el cuestionario hace referencia a los *elementos y procesos implicados en la comprensión y en su mejora*. En este sentido, entre los aspectos considerados más relevantes a la hora de comprender, las respuestas se dirigen mayoritariamente a destacar aquellos procesos implicados en la comprensión que tienen que ver con el propio alumnado. Entre ellos, los referidos a procesos implicados en la integración textual, son los que abundan entre las respuestas dadas. En este sentido, se hace alusión, entre otros, a aspectos relacionados con: (a) la comprensión del significado; (b) la comprensión del vocabulario; (c) la distinción de la información relevante; (d) la distinción de la estructura textual; (e) la capacidad para representarse mentalmente el contenido del texto; o (f) la capacidad para resumir el texto con las propias palabras, captando el contenido global y la intención del autor. Con relación a los procesos metacognitivos, se destacan de forma explícita tan sólo cuatro respuestas, que aluden a los procesos de metacognición en general, a la importancia de detenerse en los conceptos que no se comprenden (dos referencias), y a la importancia de centrarse especialmente en el contenido global de un texto y no sólo en el significado concreto de algunas palabras. Por lo que respecta a las respuestas que hacen referencia a otros aspectos no centrados en el propio alumnado, se destacan aquellos que tienen que ver con los textos, volviendo a reseñar la importancia de que el texto y el vocabulario sean

adecuados al nivel del alumnado y de que los textos sean igualmente interesante para ellos.

Entre los elementos que favorecerían la comprensión en el alumnado, podemos observar que se destacan especialmente aquellos que tienen que ver con el contexto escolar y especialmente con el proceso de enseñanza y aprendizaje, destacando aspectos metodológicos centrados en algún proceso concreto de lectura. También se alude a la metodología relacionada con los textos a utilizar, volviendo a destacar la relevancia del uso de textos adecuados, motivadores, bien estructurados y que permitan la realización de actividades lúdicas e interesantes para el alumnado. Con relación a la cuestión anterior, el profesorado destaca en otra de las preguntas de este bloque, que sería necesario intervenir sobre procesos implicados en el ajuste personal-social para mejorar la comprensión, incluyendo aspectos relacionados con la motivación y el interés de los sujetos, pero haciendo igualmente referencia a la necesidad de intervenir sobre el proceso mismo de enseñanza-aprendizaje. En cuanto al resto de procesos contemplados, más centrados en el sujeto, vuelven a considerarse los procesos implicados en la integración textual como los más relevantes, mientras que sólo una de las respuestas destaca la importancia de intervenir sobre los procesos metacognitivos para mejorar la comprensión de textos. Por otra parte, en otra de las preguntas del bloque, relacionada con la información que es necesario recoger para la evaluación de la comprensión de textos, vuelve a destacarse como elemento relevante la alusión a los procesos implicados en la integración textual. En segundo lugar, en cuanto al número de referencias realizadas, se alude a los procesos psicológicos subyacentes y a los procesos de asociación G-F, mientras que la información relacionada con los procesos implicados en la integración texto-conocimientos previos, así como a los procesos metacognitivos, reciben el menor número de referencias.

Con relación a la *relevancia otorgada por el profesorado a la colaboración con otros profesionales y con la familia del alumnado*, la mayoría de los docentes da muestras, mediante su valoración cuantitativa, de considerar muy importante la colaboración con otros profesionales a la hora de afrontar las dificultades de aprendizaje relacionadas con la lectura, obteniéndose una valoración media de 4,07 sobre 5, con diez profesores que consideran máxima su importancia. Esta valoración relativa a la importancia de la colaboración, es aún más alta cuando se hace referencia a la familia del alumnado; en este caso, la valoración media es de 4,78 sobre 5, con veintiún profesores que otorgaron el máximo valor posible. Por su parte, en lo relativo a la *consideración del lugar que debe ocupar la enseñanza y el trabajo sobre la lectura dentro del currículum*, el conjunto de profesores y profesoras participantes del estudio refleja de forma muy mayoritaria que la lectura debería ser trabajada durante todo el periodo de escolarización del alumno.

El penúltimo de los bloques de preguntas del cuestionario hace alusión a las *expectativas, la valoración y las ideas relacionadas con la atención educativa al alumnado con NEE*. En este sentido, el profesorado manifiesta en general expectativas positivas hacia las posibilidades de aprendizaje de los sujetos con dificultades, aunque en una parte importante de las respuestas se hace referencia a aquellos aspectos que serían necesarios para incrementar estas posibilidades. Entre estos, se citan por ejemplo: (a) la formación del profesorado; (b) la implicación de la familia; (c) la necesidad de apoyo social, escolar u organizativo; (d) la necesidad de trabajar en equipo; etc. Sin embargo, ante una cuestión menos general, formulada con relación a la experiencia

particular de los docentes, se hace referencia en mayor medida a las dificultades e incluso a la imposibilidad de atender adecuadamente al alumnado que presenta NEE. No obstante, siguen siendo mayoritarias las respuestas que expresan valoraciones y expectativas positivas con relación a la posibilidad de atención. En este caso, la mayoría de las respuestas reflejan una percepción positiva pero condicionada a los medios y recursos disponibles, así como al trabajo de coordinación con otros profesionales y con la familia. Por otra parte, el profesorado expresa mayoritariamente que las dificultades de aprendizaje relacionadas con la lectura se dan en cualquier circunstancia y con cualquier texto; no obstante, un tercio de las respuestas recogidas hace referencia concreta a situaciones de aprendizaje (relativas al tipo de texto y a aspectos metodológicos centrados en diversos procesos implicados en la comprensión, especialmente en procesos de ajuste personal-social) bajo las cuales las diferencias se atenúan o acentúan. Bajo nuestro punto de vista, este último tipo de respuestas, podría estar más en consonancia con procesos de evaluación e intervención que tratan de «afinar» más en la ubicación concreta de las dificultades de lectura, así como en la búsqueda de la mejor solución con relación a la mejora del proceso. De esta forma, de estas respuestas se derivarían procedimientos que habrían sido capaces de detectar en qué situaciones y bajo qué condiciones pueden experimentar algún cambio las dificultades que manifiesta el sujeto.

Con relación al resultado de las intervenciones realizadas para mejorar la lectura en el alumnado que presenta dificultades, éste es visto en general como positivo por el profesorado, aunque un análisis minucioso del contenido de sus respuestas, denota de forma mayoritaria que los avances no son suficientes o están muy condicionados por las dificultades relativas al propio alumnado o al contexto social-familiar. Por último, en cuanto al *conocimiento y trabajo con programas específicos de comprensión de textos*, la mayoría de los docentes manifestó no conocer ningún programa de estas características y, lógicamente, no haber trabajado con ellos. Por otra parte, algunos de los programas que se mencionan son en realidad fichas de actividades sobre comprensión lectora, que no constituyen un programa sistemático y global para la mejora del proceso de comprensión. La constatación de la necesidad de formación del profesorado en este campo, fue expresada sin embargo tan sólo en una de las respuestas.

Relación entre las puntuaciones de mejora y las respuestas en EIDAL

Hemos procedido a correlacionar las respuestas obtenidas en el cuestionario de expectativas e ideas relativas a las dificultades de aprendizaje de la lectura, con las puntuaciones de mejora en las pruebas de comprensión para primaria y secundaria, así como con las puntuaciones de mejora obtenidas en la escala APSL y en Cattell-2. El análisis de los resultados de estas correlaciones perseguía sobre todo establecer alguna posible relación significativa entre la mejora de los sujetos, por una parte, y determinadas expectativas, ideas o actuaciones relativas a la atención educativa del alumnado con dificultades, por otra. Lógicamente, el análisis ha sido efectuado en función de los cuestionarios cumplimentados y entregados, por lo que una de sus limitaciones principales consiste en que no recoge el sentir y el pensar de todos los docentes que han participado de una u otra forma en el estudio. De los 27 cuestionarios utilizados para confeccionar el sistema de categorías expuesto y analizado anteriormente, 15 pertenecían a profesores y profesoras con alumnado incluido en el grupo experimental. En este sentido, la matriz de datos construida para el análisis de las correlaciones, ha contemplado que algunos docentes impartían clase a más de un sujeto,

por lo que hemos establecido relación entre las distintas puntuaciones de mejora obtenidas por los diferentes sujetos y las respuestas otorgadas por su profesor. De esta forma, la n resultante para el análisis, responde mejor al número real de posibles relaciones establecidas, estando constituida finalmente por la suma de las posibles correlaciones que podían tener lugar entre respuestas del profesorado, por un lado, y puntuaciones de mejora, por otro ($N = 56$ con relación a las pruebas de comprensión; $N = 51$ para la escala APSL; y $N = 55$ para Cattell 2). Por otra parte, también debemos reseñar que no todas las correlaciones analizadas entre las puntuaciones de mejora y las respuestas del cuestionario, parecen tener igual relevancia para la investigación. En este sentido, hemos destacado tan sólo aquellas que podían generar interés desde el punto de vista de la aportación de información relevante al estudio, o de abrir alguna vía de interpretación significativa. Hemos de tener en cuenta igualmente que, en su mayor parte, se trata de ideas, expectativas y valoraciones efectuadas por el profesorado a quien en principio estaba dirigido el cuestionario, es decir, profesores y profesoras que ejercían funciones de tutoría del alumnado participante, o bien que impartían clase en áreas de contenido lingüístico; por tanto, podemos echar en falta determinadas respuestas más acordes con perspectivas más dinámicas y metacognitivas, y que podrían haber encontrado niveles de correlación positiva y significativa con las puntuaciones de mejora en caso de que todos los aplicadores del dispositivo de evaluación dinámica, integrantes del Grupo de Trabajo, hubiesen cumplimentado el cuestionario.

Al igual que hemos hecho con el análisis exploratorio de las respuestas obtenidas en los cuestionarios, nos apoyamos en la estructura de los distintos bloques contemplados en el sistema de categorías construido para mostrar el análisis correlacional efectuado.

Con relación al primer bloque de cuestiones, la puntuación de mejora obtenida en las pruebas de comprensión ECO 1 y ECOS, correlaciona muy significativamente con las respuestas del profesorado que centran las dificultades en el propio alumnado, y más concretamente en los procesos psicológicos subyacentes: *falta de atención y dificultades de memoria* ($r.0,356$ para ambas respuestas). Igualmente, la mejora correlaciona muy significativamente con las respuestas que sitúan las dificultades de aprendizaje de la lectura en *los procesos de conversión G-F* ($r.0,356$) y en *las dificultades de articulación* ($r.0,356$). Con idéntica intensidad correlaciona con las respuestas que señalan a *las dificultades del alumnado en la comprensión* (procesos de integración textual) como el factor principal que explicaría las dificultades ($r.0,356$). Sin llegar a ser significativa ($r.0,164$), la mejora también correlaciona positivamente con las respuestas que sitúan las principales dificultades en los procesos metacognitivos, y más concretamente en los *problemas relacionados con la impulsividad*. Por otra parte, la mejora correlaciona negativamente ($r.-0,294$) con las respuestas que indican que las dificultades se sitúan en *la falta de motivación/implicación/dedicación en el contexto familiar*. Este dato posee un interés especial, dado que podría llegar a explicar la ausencia de ganancias significativas precisamente en aquellos casos en los que el profesorado ha destacado las situaciones de privación socio-cultural y familiar como principal factor explicativo de las dificultades relacionadas con la lectura. En línea con lo expuesto más arriba, hallamos igualmente una correlación negativa y significativa de la puntuación de mejora con las respuestas que hacen referencia al contexto escolar como fuente de las dificultades de aprendizaje. Por ejemplo, con aquellas que hacen referencia a que *algunos procesos no se trabajan* ($r.-0,268$). Posteriormente, en el capítulo de Discusión abordaremos las posibles explicaciones de estos resultados.

Siguiendo con nuestro análisis, la puntuación de mejora en comprensión correlaciona significativamente con las respuestas que destacan el contexto social como factor explicativo de las dificultades de aprendizaje de la lectura, especialmente *la influencia de los medios de comunicación audiovisuales* (r.0,279). Igualmente correlaciona de forma muy significativa con las respuestas que indican los *factores relacionados con los textos que deben leer los alumnos, como aquellos más relevantes en la explicación de las dificultades* (r.0,361). También correlaciona positivamente, aunque sin llegar a la significatividad estadística, con los factores relacionados con el propio alumnado: *sus propias dificultades, su grado de motivación, etc.* (r.0,205). En cambio, la mejora vuelve a correlacionar negativamente con las respuestas que sitúan el contexto socio-familiar y escolar como los principales factores explicativos.

En cuanto a la mejora observada en la escala APSL, hay correlación positiva (r.0,215) con las respuestas que destacan los *factores relacionados con el propio alumnado como principales en la explicación de las dificultades*. El análisis de las respuestas dadas en otra de las cuestiones nos permite afinar más en cuanto a qué tipo de factores son los relevantes en la explicación de las dificultades; en el sentido apuntado anteriormente, y también sin llegar a tener un valor significativo estadísticamente, la mejora obtenida en la escala APSL correlaciona positivamente (r.0,238) con las respuestas que sitúan las principales dificultades en los procesos de ajuste personal-social, y más concretamente en *la falta de motivación y esfuerzo*.

Con relación a los motivos que explicarían la escasa motivación del alumnado, encontramos una relación negativa entre la mejora en comprensión y las respuestas que hacen referencia a algunos motivos relacionados con el propio alumnado; por ejemplo, con las respuestas: *el alumno tiene dificultades de comprensión* (r.-0,321); *no tienen fluidez y leer les supone un gran esfuerzo* (r.-0,301); *la lectura les resulta poco útil y significativa* (r.-0,307). Por otra parte, correlaciona positiva y significativamente con respuestas que indican motivos de índole familiar; por ejemplo: *la asociación de la lectura con una actividad sólo escolar y no significativa* (r.0,263). También con relación a la escala de ajuste, sin llegar a la significatividad, encontramos correlación entre la mejora y los motivos relacionados con el contexto familiar: *entorno familiar no interesado o no colaborador* (r.0,246), y en la misma línea, con respecto al contexto social: *contexto social no favorecedor* (r.0,217).

Con relación a las dificultades de aprendizaje concretadas en un alumno, encontramos correlación muy significativa entre la puntuación de mejora en comprensión y las referencias del profesorado a *la falta de concentración y atención* (r.0,356). La mejora también correlaciona muy significativamente con las respuestas que hacen referencia a los procesos metacognitivos: *impulsividad* (r.0,350) y a los procesos de ajuste personal-social: *falta de motivación* (r.0,356). En cambio, correlaciona negativamente con las respuestas que destacan las dificultades relacionadas con los procesos de integración textual: *problemas para resumir y extraer lo relevante*, y con otros procesos metacognitivos: *no utiliza procesos de autorregulación o no reflexiona* (r.-0,301 y r.-0,301 respectivamente).

La cuestión que aborda el origen de las dificultades de aprendizaje del alumno al que se hacía referencia anteriormente, o los elementos que mantienen estas dificultades, ofrecía también una correlación negativa y muy significativa entre la mejora en comprensión y las respuestas que destacan la relevancia del contexto familiar y social próximo, y más

concretamente: *el escaso apoyo familiar y del contexto* (r.-0,349). En cambio, hallamos correlación positiva, aunque no llega a ser significativa, entre la mejora y las respuestas que sitúan el origen de las dificultades o los factores que las mantienen, en los procesos relacionados con el contexto escolar, y más concretamente en el proceso de enseñanza-aprendizaje, por ejemplo: *no haber sido motivados adecuadamente en la escuela* (r.0,151); o *pensar que a partir de un cierto nivel, la comprensión lectora ya está conseguida* (r.0,151).

El segundo bloque de cuestiones aborda la metodología en la enseñanza/facilitación y evaluación de los procesos de lectura y comprensión. Con relación a la cuestión que abordaba una metodología basada en la discusión y la confrontación de ideas, conocimientos, experiencias relacionadas con la lectura, o estrategias que son usadas para comprender, no hallamos ninguna correlación significativa; tan sólo las respuestas que destacan la relevancia de que los alumnos *comenten con el profesorado el libro que están leyendo, señalando su propia opinión, dificultades encontradas, etc., para lograr un razonamiento acerca de qué les gusta y por qué*, correlaciona positivamente con la mejora, siendo el valor más alto que encontramos: r.0,218. Con relación a cómo se trabajarían en el aula los aspectos relativos a la comprensión considerados relevantes por el profesorado, encontramos una relación muy significativa entre la puntuación de mejora y las respuestas que destacan *la elección/adaptación de los textos y temas en función del alumnado, sus capacidades e intereses* (r.0,429). En cambio, *la modelación del uso de estrategias para trabajar los diferentes procesos y la enseñanza/facilitación de estrategias, mediando para facilitar la interiorización de estrategias relacionadas con los diversos procesos*, correlaciona negativamente con la mejora (r.-0,268 en ambos casos). Por otra parte, encontramos correlación positiva de la mejora, aunque sin llegar a ser significativa, con otros aspectos relacionados con la facilitación del aprendizaje estratégico, por ejemplo: *la enseñanza de estrategias, trabajando la estructura textual en relación con los diferentes tipos de textos, y las claves textuales* (r.0,106); y *el análisis/diálogo comprensivo sobre el desarrollo de los textos, intentando desarrollar la comprensión lectora y la expresión, fomentando la relación con los conocimientos y experiencias previas* (r.0,218). También correlaciona de forma muy significativa la mejora en comprensión con las acciones encaminadas a la creación de un clima adecuado para la lectura, destacando *la motivación previa a la lectura* (r.0,356).

No hemos encontrado, por otra parte, relaciones significativas entre las respuestas que destacaron los elementos más relevantes a la hora de la enseñanza de la lectura, y las puntuaciones de mejora en comprensión. Tan sólo con relación a la escala de ajuste personal-social observamos que varias respuestas obtienen correlaciones positivas, por ejemplo: *la presentación atractiva y motivadora de los textos y de los temas* correlaciona positivamente (r.0,251) con la mejora de los procesos de ajuste; igualmente *la metodología y las actividades adecuadas al nivel y las necesidades del alumnado* (r.0,200); y *la motivación previa a la lectura* (r.0,200).

En cuanto a las actuaciones que se han llevado a cabo para paliar las dificultades de aprendizaje, también encontramos correlación muy significativa de la puntuación de mejora en comprensión con las medidas que tratan de incidir en los procesos de ajuste personal-social, y más concretamente en *la creación de un clima de aula adecuado* (r.0,356). También obtienen niveles de correlación altos, aunque sin llegar a la significatividad, las respuestas que hacen referencia al *uso de textos adecuados y amenos, relacionados con los intereses del alumnado* (r.0,255). En cuanto a la escala de

ajuste personal-social, *proponer un trabajo individualizado y adaptado*, correlaciona positivamente con la mejora (r.0,200), al igual que *dedicar más tiempo al alumnado con dificultades y a la lectura, trabajando con ellos* (r.0,200).

Con relación a los elementos necesarios para incrementar la motivación, hemos encontrado correlación positiva y significativa entre la mejora y una respuesta que alude al proceso de enseñanza-aprendizaje, más concretamente a *la realización de actividades previas a la lectura (motivación previa) y también posteriores (comentarios, invención de nuevos textos con el mismo tema, variación del final...)* (r.0,322). También hallamos correlación positiva con otras respuestas que hacen referencia a: *buscar textos atractivos que conecten con sus intereses y presentación atractiva de los textos* (r.0,182); y *plantear la lectura en voz alta como un premio* (r.0,218 con la mejora en comprensión y r.0,170 con la mejora en APSL). Con relación a la escala de ajuste, la puntuación de mejora correlaciona positivamente con respuestas que destacan la importancia de *desarrollar estrategias más atractivas y motivadoras, y transmitir entusiasmo por la lectura* (r.0,200). Por otra parte, no encontramos correlaciones significativas de la puntuación de mejora con relación a la necesidad de adecuación de la metodología en el aprendizaje de la lectura; el valor más alto se da con la respuesta que hace referencia a que un mismo método es aplicable en general, destacando la necesidad de detectar necesidades y realizar actividades adecuadas (r.0,115).

Con relación a la evaluación de la comprensión, hallamos correlación negativa y significativa entre la mejora en comprensión y la *evaluación mediante lectura y preguntas posteriores (orales o escritas)* (r.-0,306). Igualmente, *la observación directa de la lectura y el diálogo para construir la comprensión, realizando preguntas individuales por escrito*, mantiene correlación negativa con la mejora (r.-0,268). En cambio, aunque sin llegar a ser significativa, encontramos correlación positiva con las respuestas que hacen referencia a una *evaluación continuada (casi a diario) mediante la realización de preguntas sobre un personaje, su opinión, relación del texto con experiencias personales* (r.0,218).

El tercer bloque de cuestiones hace referencia a los elementos y procesos implicados en la comprensión y en su mejora; es decir, aquellos aspectos que son considerados por el profesorado como más relevantes a la hora de comprender un texto. Obtenemos una correlación negativa y significativa (r.-0,268) entre la mejora en las pruebas de comprensión y las respuestas que destacan la relevancia de los *procesos de asociación G-F en los inicios del aprendizaje de la lectura*. También correlacionan negativamente y de forma significativa (r.-0,268) con la puntuación de mejora, las respuestas que destacan la relevancia de los *aspectos relacionados con los procesos de integración texto-conocimientos*, así como aquellos relacionados con los *procesos metacognitivos*. Tal y como apuntábamos anteriormente, podría estar influyendo decisivamente en los resultados de este análisis, la situación de desventaja socio-cultural de parte del alumnado de primaria que participó del grupo experimental. La mayoría de estos sujetos no obtuvieron mejoras posttest en comprensión, por lo que las respuestas destacadas por sus profesores mantienen habitualmente una correlación negativa con la puntuación de mejora. Por otra parte, aunque a priori existía la posibilidad de establecer una comparación dentro de uno de los grupos de alumnos pertenecientes a un centro de especiales dificultades, comparación que tendría lugar entre los participantes del grupo experimental y del grupo control, lo que hubiera sido importante para comprobar en qué medida la similitud o las diferencias en las respuestas se correspondían con la obtención

de distintos resultados; el conocimiento posterior proporcionado por los profesores colaboradores acerca de la formación de los distintos agrupamientos en el centro, restaba en gran medida interés a este análisis. En este sentido, una de las aplicadoras informaba sobre los distintos niveles de partida relativos a la competencia curricular de los sujetos que integraban los grupos control y experimental en este centro. De esta forma, el grupo experimental estaba formado por alumnos y alumnas que participaban de un agrupamiento flexible caracterizado por un nivel de competencia en lectura significativamente más bajo que el de los sujetos que formaban parte del grupo control. Por otra parte, encontramos que correlacionan positiva y significativamente con la mejora en comprensión ($r=0,407$), la consideración de los *aspectos concernientes a la adecuación de los textos y del vocabulario al nivel del alumno*. Igualmente, el hecho de que *los textos propuestos sean interesantes para el alumnado* ($r=0,356$). Por su parte, aunque sin llegar a la significatividad, también mantienen una correlación positiva, *la relevancia y la adecuación de la estructura textual en los textos propuestos* ($r=0,194$); *la concentración en la lectura* ($r=0,164$) y *la sintaxis* ($r=0,164$).

En cuanto a los elementos que se creen necesarios para favorecer la comprensión de textos, no encontramos ninguna relación significativa con la mejora, obteniendo los valores más altos de correlación las respuestas que hacen referencia a elementos relacionados con el contexto escolar (procesos de enseñanza-aprendizaje), y más concretamente: *la ampliación de los conocimientos del alumnado —teóricos y prácticos— en todas las materias*, con lo que se trataría de elementos relacionados a su vez con los procesos de integración texto-conocimientos ($r=0,164$); y *la realización de actividades de comprensión* ($r=0,151$). Sin embargo, con relación a los elementos sobre los que habría que actuar para mejorar la comprensión lectora, hallamos una relación muy significativa entre la puntuación de mejora y las respuestas que destacan *el trabajo sobre procesos de atención y concentración durante la lectura* ($r=0,416$). Sin llegar a ser significativa, *el trabajo sobre comprensión* ($r=0,164$), y la actuación sobre los procesos de ajuste personal-social, en concreto sobre *la motivación y los intereses del alumnado*, también correlacionan positivamente con la mejora ($r=0,203$).

En cuanto a la información que se cree relevante recoger cuando se evalúa la comprensión, hallamos correlación significativa ($r=0,279$) con las respuestas que señalan algunos aspectos relacionados con los procesos implicados en la asociación G-F, por ejemplo *la velocidad lectora*. También correlacionan positivamente las respuestas que hacen referencia a los procesos de integración textual, más concretamente: *la comprensión* ($r=0,198$); y *el conocimiento y adquisición de vocabulario general y específico de las distintas áreas* ($r=0,151$). Por su parte, las respuestas que destacan la información relacionada con los procesos metacognitivos, también mantienen correlación positiva con la mejora: *cómo lee, dónde encuentra dificultades y qué hace ante esas dificultades, si vuelve atrás, si se muestra reflexivo o impulsivo...* ($r=0,164$). Por último, también de forma no significativa correlacionan algunas de las respuestas que señalan aspectos relacionados con los procesos de ajuste personal-social: *si se muestra interesado o abandona ante las dificultades* ($r=0,164$). Dentro de estos últimos, también se incluyen algunas respuestas relacionadas con los procesos de enseñanza y aprendizaje: *observar cómo lee y qué tipo de errores comete en su lectura* ($r=0,164$).

Por lo que concierne a las cuestiones que integran el cuarto bloque de nuestro sistema de categorías, aquel que hace referencia a la colaboración con otros profesionales y con las familias del alumnado, encontramos niveles de correlación que en ningún caso

llegan a ser significativos. En cuanto a la colaboración con otros profesionales, la puntuación de mejora en las pruebas de comprensión obtiene el nivel de correlación más alto con las respuestas que otorgaron una puntuación 3 en la escala propuesta ($r=0,218$), mientras que con relación a la mejora en la escala de ajuste, es la puntuación 5 la que mantiene la correlación más alta ($r=0,142$). Por su parte, la valoración 4 otorgada a la colaboración con la familia de los sujetos, presenta la correlación más alta con la mejora en comprensión ($r=0,119$), aunque no llega tampoco a ser significativa.

En cuanto al papel que debe jugar la lectura dentro del currículum, cuestión central en el quinto de los bloques de nuestro sistema de categorías, no hallamos una relación significativa de la mejora en comprensión, o en la escala de ajuste personal-social, con las distintas opciones relativas al periodo en que debe ser trabajada la lectura. En este sentido, ya hemos podido comprobar como, aunque con matices, la gran mayoría de los docentes expresaron su convencimiento de que la lectura debía ser objeto de trabajo al menos durante todo el periodo de escolarización obligatoria del alumno. En general, podríamos deducir a tenor de los resultados obtenidos, que las opiniones sobre el periodo de tiempo que debe ser trabajada la lectura en la escuela, así como sobre la organización de la enseñanza de la misma, parece tener poca incidencia en la actuación práctica con relación a la enseñanza misma de la lectura y la comprensión.

Con relación al bloque de cuestiones que hacen referencia de forma explícita a las ideas, expectativas y valoración sobre el aprendizaje del alumnado que presenta necesidades educativas especiales, encontramos correlaciones positivas, aunque sin llegar a ser significativas, con las respuestas que expresan expectativas positivas: *cualquier avance es importante, sobre todo en una actividad tan relevante como la lectura* ($r=0,218$); o *me parece fundamental el aprendizaje de la lectura y la escritura para eliminar barreras* ($r=0,218$). Encontramos, así mismo, una correlación muy significativa ($r=0,356$) con las respuestas que señalan *la necesidad de adaptar los textos a sus necesidades e intereses para incrementar las posibilidades del alumnado con NEE*. En cuanto a la mejora en la escala de ajuste, hallamos correlación positiva con las respuestas que destacan *la motivación como factor fundamental para el incremento de las posibilidades de este alumnado* ($r=0,200$). Un aspecto que resulta de especial interés, aunque no esté avalado por la obtención de niveles significativos, es que las respuestas que mostraron expectativas centradas en las limitaciones del alumnado con NEE y que hicieron referencia explícita a factores intrínsecos del propio alumnado, como su grado de déficit, para explicar las posibilidades de aprendizaje, correlacionaron negativamente con la mejora en comprensión ($r=-0,205$), así como también con la mejora en la escala de ajuste ($r=-0,215$) y en la prueba de CI ($r=-0,060$).

En cuanto a si pueden ser atendidas adecuadamente las necesidades educativas especiales relacionadas con la lectura, encontramos correlación significativa con respuestas que inciden en la *dificultad de atención debido al elevado número de alumnos, imposibilitando una atención educativa adecuada* ($r=0,287$). Hallamos también correlación positiva y muy significativa ($r=0,356$) con las respuestas que señalan la *dificultad de atención adecuada debido al bajo grado de motivación del alumnado*; y con las que inciden en los *problemas de disciplina como causa de las dificultades para poder atender bien al alumnado con NEE* ($r=0,356$). Por lo que respecta a las respuestas que expresan expectativas positivas, éstas también destacaban condicionantes necesarios para que la atención pudiera ser efectiva. En este sentido, encontramos algunas correlaciones positivas con estas respuestas que, sin embargo, no

llegan a ser significativas ($r=0,151$ en todos los casos): *depende de los recursos de los centros; la coordinación entre profesores y los medios disponibles debe conjugarse para ser efectiva; se ha intentado volcar medios materiales y humanos en la atención.* Por otra parte, hallamos correlación positiva y muy significativa entre la mejora en comprensión y las respuestas que especifican momentos o situaciones concretas en las que se manifiestan de forma especial las dificultades en la lectura, por ejemplo: *las dificultades disminuyen si trabaja individualmente bajo mi supervisión* ($r=0,356$); y, aunque sin ser significativa: *las dificultades están más relacionadas con la lectura y la escritura que con las matemáticas, presenta dificultades en el orden y presentación de sus trabajos* ($r=0,218$). Con relación a los resultados de las actuaciones emprendidas para mejorar la lectura en los sujetos con dificultades de aprendizaje, encontramos correlación positiva entre la puntuación de mejora en comprensión y las respuestas que hacen referencia a resultados positivos aunque no suficientes, o que estuvieron condicionados por las dificultades, por ejemplo: *depende de que se den circunstancias favorecedoras (ausencia de problemas de conducta) para que los resultados sean positivos* ($r=0,218$); y, con idéntico valor de correlación: *han mejorado, pero no lo suficiente, muy lentamente.*

Por último, con relación al bloque de cuestiones que hace referencia al conocimiento y al trabajo con programas específicos para la mejora de la comprensión, el análisis de las correlaciones muestra una relación positiva, aunque no significativa, entre la mejora en comprensión y las respuestas que destacan el *desconocimiento de programas y la necesidad de formación del profesorado en este tema* ($r=0,218$); estas mismas respuestas correlacionan también positivamente con la mejora en la escala de ajuste ($r=0,170$). En cuanto a la valoración sobre los resultados de haber trabajado con programas de comprensión, encontramos correlación muy significativa ($r=0,356$) con respuestas que hacen referencia a que *los resultados no fueron tan positivos como se esperaba, y a que no se pudieron trabajar bien debido a la existencia de problemas de conducta.*

Aún una reflexión más, relativa al análisis efectuado. En línea con algunas de las limitaciones que apuntábamos al comienzo, algunos de los resultados que hemos expuesto podrían estar reflejando, más allá de la idoneidad de determinadas ideas o actuaciones, las diferencias observadas en los resultados obtenidos en general por los grupos de primaria y secundaria. En este sentido, las respuestas dadas por el profesorado de secundaria serían las que habrían primado, obteniendo correlaciones positivas y significativas con las puntuaciones de mejora, ya que fue en esta etapa donde se produjeron mejoras significativas en comprensión para el grupo experimental completo, así como para la mayoría de los sujetos.

Análisis exploratorio del cuestionario EIDAL-MP

En cuanto al cuestionario EIDAL-MP (Anexo XII), el número de familias participantes fue de 29. En este caso, todas las madres o los padres que cumplimentaron el cuestionario, lo eran de sujetos que participaban como integrantes del grupo experimental ($N = 15$) o del grupo control ($N = 14$). Por otra parte, al igual que ya hicimos con relación a la información obtenida en el cuestionario dirigido al profesorado, hemos organizado el sistema de categorías con relación a las dimensiones contempladas en el cuestionario en función del contenido de las cuestiones que plantea; en este caso, estas dimensiones son 10. El sistema de categorías construido, que recogemos en el volumen de anexos (Anexo XII-A), junto con el número de referencias

realizadas por el conjunto de las madres y padres participantes en el estudio, puede ofrecernos, al igual que en el caso del cuestionario dirigido al profesorado, información relevante para la contextualización del proceso de aplicación del dispositivo EDPL, permitiéndonos, en este caso, «visualizar» más allá de lo estrictamente acontecido en las aulas. De este modo, podemos llegar a intuir cómo tienen lugar los procesos educativos relacionados con el aprendizaje y, más concretamente, con el lenguaje escrito, en los ambientes social y familiar del alumnado que participó en la investigación. Con esta información será posible llegar a establecer algunas relaciones entre el proceso de aprendizaje llevado a cabo durante la fase de aplicación —y también entre los resultados obtenidos—, y las expectativas, ideas, valoración o actuaciones concretas que tienen lugar habitualmente en el contexto social y familiar. A continuación realizamos un recorrido a través de los distintos bloques de los que consta el sistema de categorías, destacando algunos puntos que consideramos especialmente relevantes con relación a la información obtenida.

El primer bloque de cuestiones hace referencia a diversos supuestos prácticos relacionados con las dificultades de aprendizaje de la lectura. En primer lugar, del análisis de las respuestas obtenidas en las cuestiones pertenecientes a este primer bloque, podemos desprender que las familias expresan una notable preocupación por los sentimientos de rechazo a la lectura de sus hijos e hijas. Cuando se les pregunta cómo actuarían en el supuesto de que su hijo manifestara este rechazo, sus respuestas denotan actuaciones dirigidas sobre todo a aspectos de índole socio-personal, como la animación y la motivación a la lectura. Por otra parte, la ayuda que ofrecerían en caso de presencia de dificultades de aprendizaje, se centraría en la intervención directa, mediante la dedicación de tiempo de ayuda concreta en las tareas escolares, así como en la búsqueda de apoyos externos para que orientaran la actuación. En ambos casos sólo hay una referencia explícita a la necesidad de consultar con el profesor de su hijo).

Ante la situación hipotética planteada en una de las cuestiones, en la que se habla de niños o niñas especialmente difíciles a la hora de prestarles ayuda en las actividades escolares, podemos observar que los padres dan un mayor número de respuestas que expresan intentos de ayuda directa en detrimento de aquellas que expresan actuaciones que contemplan opciones de ayuda indirecta o externa. Entre estas últimas, la búsqueda de ayuda en un especialista es la opción más nombrada. Por otra parte, la actuación de los padres y madres ante la comisión de continuos errores de sus hijos relacionados con la lectura, parece centrarse en la ayuda/intervención directa, primando de nuevo la animación, la explicación y la corrección de las tareas. Nuevamente, la búsqueda de apoyo o de colaboración con el profesorado, es recogida sólo en las respuestas de dos familias.

Ante la cuestión formulada referente a la petición de ayuda de los hijos para la realización de las tareas, algo más de la mitad de las madres y padres expresan que estos sí solicitan ayuda. Entre estas respuestas, dos hacen referencia a la escasa ayuda que pueden ofrecerles en respuesta a sus demandas, debido principalmente a sus propias dificultades o a la falta de conocimientos. El resto de respuestas recogidas expresa que sus hijos no solicitan ayuda o lo hacen poco. En dos de estos casos, en los que no hay petición de ayuda por parte de los hijos, los padres expresan su deseo de que ésta tuviera lugar, manifestando interesarse por las tareas escolares aunque no se les pida ayuda expresamente.

En las situaciones en las que los hijos comentan con sus padres lo que han hecho, o les plantean dudas relacionadas con los contenidos que han debido leer y trabajar en la escuela, la mayoría de los padres expresan en sus respuestas una actuación centrada en la atención y la escucha, por una parte, y en la ayuda directa a sus hijos, por otra. En cuanto a los refuerzos o premios que son otorgados por parte de las familias con relación al trabajo escolar, existe una cierta igualdad entre el número de referencias que aluden, por un lado, al establecimiento de refuerzos ante el trabajo y la actividad desarrollada diariamente, y aquellas que expresan, por otro, aspectos relacionados con las notas escolares o con las evaluaciones finales. Así mismo, igual número de referencias obtiene la ausencia de reforzamiento por parte de las familias ante el trabajo escolar. En línea con lo anterior, con relación al tipo de reforzamiento utilizado, parece primar el reforzamiento de carácter social, aunque esto no haya sido explicitado por parte de todas las familias. Con relación a los reforzamientos otorgados en función de las actividades desarrolladas diariamente, sí parece ser relevante para los padres, a tenor de algunas de sus respuestas, la consideración de aspectos relacionados con procesos de ajuste personal-social, como por ejemplo: (a) la autonomía mostrada en las tareas a realizar, (b) el interés manifestado, y (c) la percepción de que la actividad ha supuesto un esfuerzo para el sujeto.

Con relación a la implicación de la familia en la atención educativa relativa a las dificultades de aprendizaje, las respuestas analizadas expresan de forma muy mayoritaria que los padres pueden y deben colaborar en el desarrollo escolar, contribuyendo a la mejora del aprendizaje de la lectura. En este sentido, los padres parecen hacer referencia a la colaboración con el profesorado, manifestando que *entre todos es mejor* y que *todos podemos poner algo de nuestra parte* para conseguir mejorar las habilidades relacionadas con la lectura. Tan solo una respuesta hace alusión a que *la lectura es cosa de la escuela y de los maestros*. Cuando se les pide que concreten la forma en la que podrían influir en el aprendizaje y desarrollo de la lectura, todos coinciden en que efectivamente pueden ejercer cierta influencia, aunque difieren en las acciones que emprenderían. En este sentido, la mayoría expresa en sus respuestas que esta influencia la ejercerían, bien mediante acciones que pudieran ser entendidas como modelos a seguir por sus hijos, o bien mediante ayuda directa y estimulación/animación a la lectura. Por otra parte, un número menor de respuestas expresaban que esta influencia podría ser ejercida mediante acciones que implicaban imposición y obligación, mientras que, por último, otras no parecían tener claro cómo se podría ejercer.

En cuanto a las expectativas y deseos relacionados con los estudios que podrían realizar los hijos, las familias expresan expectativas positivas en algo más de un 60% de las respuestas, mientras que el resto de respuestas expresan expectativas negativas o desconocimiento del deseo de sus hijos. Cuando se les pregunta qué desean que sus hijos sean de mayores, la mayoría de las referencias acentúan la importancia de que la elección sea tomada por el propio hijo y de que ésta contribuya a su felicidad. También podemos observar que se hace un especial énfasis en aspectos de desarrollo personal-social, con independencia de la elección tomada (autonomía, responsabilidad, ser buena persona, cultura, educación, preparación...).

Por otra parte, la información referente a la disposición de elementos materiales relacionados con la lectura, revela algunos aspectos interesantes. En primer lugar, en algunas de las respuestas se alude a que los libros de los que se dispone no son usados

por los hijos; en cambio, algunas referencias dejan entrever un conocimiento explícito sobre el material de que se dispone, haciendo referencia a que se trata de libros infantiles o juveniles. De otras respuestas, sin embargo, parece desprenderse un cierto desconocimiento del tipo de libros que se tiene o de su adecuación con relación a la edad o los intereses de los hijos. En cuanto a la experiencia o el contacto que mantiene la familia con la lectura, las respuestas revelan que más de un 30% de los padres que cumplieron el cuestionario, no tienen contacto alguno con la lectura. A pesar de que, en concreto, este tipo de cuestiones podrían haber causado un elevado número de respuestas influidas de alguna u otra forma por la «deseabilidad social», hemos observado un alto y significativo número de referencias alusivas a la falta de experiencias relacionadas con la lectura.

Las ideas expresadas con relación a la participación conjunta y a la inclusión de los hijos en actividades relacionadas con la lectura que se está realizando o con el programa televisivo que se está viendo, denotan mayoritariamente respuestas positivas y favorecedoras de la participación y la inclusión del sujeto en la actividad; así mismo, las respuestas relativas a la consideración de idoneidad de la participación conjunta en actividades concretas de lectura, revelan un acuerdo mayoritario (más del 80%) con respecto a los aspectos positivos de estas actuaciones. Por otra parte, el resto de las respuestas hacen énfasis en las dificultades para llevar a cabo este tipo de prácticas, bien centradas en la falta de interés manifestado por los propios hijos o en la ausencia de tiempo.

Con respecto a la percepción que tienen las familias acerca de los factores que influyen en las dificultades de aprendizaje de los propios hijos, las respuestas coinciden con las de los profesores en señalar los aspectos relacionados con el propio alumnado como aquellos más relevantes en la explicación de las dificultades. Entre estos últimos, los aspectos que hacen referencia a los procesos de ajuste personal-social son los más nombrados (falta de interés y de motivación, poca dedicación, falta de esfuerzo...). Al igual que ya ocurriera en las respuestas dadas por el profesorado, son los aspectos de índole personal-social y, en segundo lugar, los relacionados con los procesos psicológicos subyacentes, los que estarían detrás de las dificultades de aprendizaje de la lectura. En menor medida aparecen los aspectos relacionados con el contexto familiar, escolar y social del alumno. Coinciden además en el análisis derivado de las respuestas de los docentes en que cuando hacen referencia al contexto escolar, destacan la importancia del interés que deben despertar los textos y actividades relacionadas con la lectura, así como en la competencia que suponen los medios audiovisuales como factor social.

Con relación a la concepción de la lectura y a su relevancia social, los padres le otorgan especial importancia como elemento válido y fundamental para el desarrollo social del sujeto (*es básico para desenvolverse en la vida*) y también para el desarrollo de cualquier aprendizaje que deba realizar. Consideran mayoritariamente que *es la base de un buen aprendizaje*. Igualmente, con respecto a la otra cuestión relativa a la concepción de la lectura, y que hacía referencia explícita al significado y sentido de la misma, es decir, para qué sirve leer, las familias destacan casi por igual elementos relacionados con la validez de la lectura para desenvolverse y defenderse en la vida, por una parte, y elementos relativos a la validez de la lectura para el aprendizaje y el enriquecimiento personal, por otra.

Las respuestas de las familias relacionadas con las expectativas sobre las posibilidades de aprendizaje del alumnado con NEE que muestra dificultades en el inicio del aprendizaje de la lectura, denotan expectativas positivas en la mayoría de los casos. Se hace referencia a elementos como la paciencia y la ayuda necesaria para que pueda tener lugar el aprendizaje, al empeño, a la ayuda especializada y al apoyo del profesorado y de los padres. Por último, con relación a la relevancia otorgada a la colaboración con el profesorado, las madres y los padres expresan de forma mayoritaria la importancia de esta colaboración, con una valoración media de 4,54 sobre 5 y un total de 19 familias que otorgaron la máxima puntuación.

Relación entre las puntuaciones de mejora y las respuestas en EIDAL-MP

Al igual que hicimos con relación al profesorado, hemos tratado de establecer algunas correlaciones entre las respuestas obtenidas en el cuestionario para madres y padres sobre expectativas e ideas acerca de las dificultades de aprendizaje de la lectura, y las puntuaciones de mejora obtenidas en cada una de las distintas pruebas criterio utilizadas: las pruebas de comprensión de textos para primaria (ECO 1) y secundaria (ECOS), la escala de ajuste personal-social APSL y la prueba de inteligencia general Cattell-2. El análisis de estas correlaciones se ha dirigido especialmente a descubrir posibles relaciones significativas entre la mejora obtenida por el alumnado que participó como grupo experimental, por una parte, y determinadas expectativas, ideas, valoraciones o actuaciones relativas a la atención educativa familiar de estos sujetos, por otra. Como ya apuntamos con relación a las respuestas del profesorado, el análisis ha sido efectuado en función de los cuestionarios cumplimentados y entregados por las familias, por lo que comparte básicamente la limitación de no recoger de forma completa el sentir y el pensar de todos los padres y madres del alumnado que ha participado en la investigación. Recordemos que el número de cuestionarios recogidos fue de 29, de los cuales 15 pertenecían a familias de alumnos y alumnas incluidos en el GE (N = 15 con relación a las pruebas de comprensión; N = 14 para APSL; y N = 15 para Cattell 2). Hemos obtenido un gran número de correlaciones; sin embargo, la exposición de la mayoría de ellas no nos ha parecido relevante en función de los objetivos del estudio. Hemos procedido, por tanto, a destacar sólo aquellas que pudieran resultar de interés con relación a otros resultados obtenidos en nuestra investigación, o a alguna de las hipótesis formuladas en la misma. Para la exposición de estos resultados, nos apoyaremos también en la estructura de los distintos bloques de cuestiones contemplados en el sistema de categorías construido.

Con relación al primer bloque de cuestiones, que contempla una serie de supuestos relacionados con las dificultades de aprendizaje, ante los cuales se pide a los padres que expongan cómo actúan habitualmente, o actuarían en el caso presentado, encontramos algunos resultados que pueden resultar de interés al relacionar las respuestas de las familias ante los diversos supuestos planteados. En primer lugar, hallamos una correlación significativa ($r=0,518$) entre la mejora en comprensión y las respuestas que hacen referencia a una actuación centrada en la animación/motivación a la lectura: *haría lo posible para que le llamara la atención y le interesara*. Esta respuesta también correlaciona positivamente, aunque no de forma significativa, con la mejora en APSL ($r=0,258$). Otras respuestas, igualmente centradas en la animación/motivación a la lectura, correlacionan también positivamente, sin llegar a ser significativas, con la mejora en comprensión: *reforzándole positivamente en su trabajo para darle ánimos* ($r=0,196$); *ayudándole con lecturas que le gustasen y le estimulen* ($r=0,196$); y con la

mejora en APSL: *le ayudaría en todo lo que pudiera* (0,507). Por su parte, la actuación centrada en la obligación y el castigo, correlaciona negativamente con la mejora en APSL (r.-0,515).

Con relación al modo en que se prestaría ayuda, encontramos correlación positiva y significativa entre la mejora en comprensión y las respuestas que indican una ayuda centrada en la intervención directa: *explicarle las cosas que no entienda* (r.0,518). Otras respuestas, también centradas en la intervención directa, correlacionan positivamente, sin llegar a la significatividad, con la mejora en la escala de ajuste: *ponerlo todos los días a leer* (r.0,398). Por su parte, las respuestas centradas en la animación/motivación a la lectura, obtienen también correlaciones positivas con la mejora en APSL: *le diría que dedicara más tiempo a la lectura* (r.0,308), aunque tampoco llegan a ser significativas estadísticamente.

Encontramos una relación positiva, aunque no significativa, entre la puntuación de mejora en la escala de ajuste y algunas respuestas que destacan actuaciones que expresan intentos de ayuda directa, si bien es cierto que en ambos casos las respuestas dejan entrever una opción por favorecer un trabajo autónomo del sujeto: *dejaría que fuera un poco independiente; sólo se ayudaría en caso de que me necesitara* (r.0,507); *vería las tareas ya terminadas, para corregirlas* (r.0,507). Otras respuestas que expresan también actuaciones centradas en intentos de ayuda directa, también obtienen correlaciones positivas sin llegar a la significatividad, con la mejora en comprensión: *hablar mucho con él, hacerle comprender que es mejor que alguien esté a su lado* (r.0,288); *le diría que si me necesita, puede pedirme ayuda* (r.0,368).

Con relación a la actuación de los padres en el supuesto de que los hijos cometan continuos errores en las actividades de lectura, hallamos correlación positiva, aunque no significativa, entre la mejora en comprensión y respuestas que expresan actuaciones centradas en la ayuda/intervención directa, especialmente dirigidas a la animación, la explicación o la corrección: *ayudarle a que se esfuerce un poco más y a comprender las cosas* (r.0,368); *le diría lo que yo pienso que es correcto, pero sin decirle la respuesta, para que piense* (r.0,368). Por otra parte, las actuaciones que no expresaban intentos de ayuda directa, o que estaban centradas en la búsqueda de ayuda externa, no obtuvieron correlaciones positivas con las puntuaciones de mejora. En varios casos, este tipo de respuestas, que de una u otra forma expresaban actuaciones centradas en la búsqueda de soluciones fuera del ámbito familiar, hallaron correlación negativa, a veces significativa, con la mejora en comprensión, así como en la escala de ajuste y también en la prueba de CI.

En cuanto a la actitud de los padres ante situaciones en las que los hijos pretenden opinar sobre importantes asuntos familiares, no se encontraron relaciones significativas entre las puntuaciones de mejora y las distintas respuestas. Los datos más relevantes muestran relaciones positivas en respuestas que indican una actuación de escucha e inclusión, que permite la participación: *escucharle para saber qué opina y cómo se siente* (r.0,346); *le hago ver que las cosas no son tan fáciles, sino que hay que sopesar porque todo tiene sus "pros" y sus "contras"* (r.0,196). Las opciones que se centraron en la exclusión y en la no escucha de las opiniones, no permitiendo la participación, obtuvieron correlaciones negativas, aunque no significativas, con la mejora en comprensión (r.-0,087) y en CI (r.-0,455): *pidiéndole que se mantenga al margen*.

Encontramos una relación positiva aunque nuevamente sin llegar a la significatividad estadística, entre las respuestas que expresan ausencia de petición de ayuda por parte de los hijos, y la mejora en comprensión y APSL. Estas respuestas tienen la peculiaridad de que muestran una actitud favorable al apoyo por parte de las familias: *no [solicita ayuda], aunque yo se las miro [las actividades] para ver si están hechas* (r.0,507 con la mejora en la escala de ajuste); *no [solicita ayuda], aunque me gustaría, para ayudarle a comprender* (r.0,368 con la mejora en comprensión). En general, estas respuestas parecen mostrar una menor necesidad de ayuda por parte de los hijos; sin embargo, también muestran que en caso de necesitar la ayuda, los padres se muestran disponibles. Esto nos lleva a constatar el hecho de que las respuestas que indican claras peticiones de ayuda por parte de los hijos, señalando el modo en que se presta este apoyo, podrían estar indicando una mayor presencia de dificultades en estos sujetos y, por tanto, la obtención de un menor grado de mejoras. De otro lado, aquellas respuestas que parecen mostrar situaciones en las que se requiere menos ayuda, y que, sin embargo, muestran disponibilidad de los padres para el apoyo en caso de ser solicitada, podrían relacionarse con la obtención de mejores puntuaciones.

No hemos encontrado relaciones significativas entre las puntuaciones de mejora y las respuestas que hacen referencia a situaciones en las que los hijos pueden comentar a sus padres hechos o actividades escolares, solicitando ayuda o planteando dudas. Entre los datos más relevantes, hallamos una relación positiva con actuaciones centradas en la atención/escucha y la ayuda directa: *ayudándole y explicándole lo que no entiende* (r.0,214 con la mejora en comprensión; y r.0,148 con la mejora en APSL). Sin embargo, las actuaciones centradas en la atención y la escucha, acompañadas de búsqueda de ayuda externa correlacionaron negativamente con la mejora en comprensión (r.-0,280) y APSL (r.-0,141): *le aconsejaría que preguntara al profesor*. Por su parte, una respuesta que alude a la ausencia de comentarios relevantes por parte de los hijos, pero que sin embargo revela el interés de los padres: *me gustaría que me preguntara y le ayudaría*, también obtiene correlación positiva con la mejora en comprensión (r.0,368).

Hallamos una relación positiva, aunque no significativa, entre la mejora en la escala de ajuste personal-social y las respuestas que indican el establecimiento de reforzamientos o premios relacionados con la realización de las actividades escolares, cuando se piensa que *el trabajo le ha costado esfuerzo y ha intentado hacerlo bien* (r.0,398). También obtenemos una relación positiva con el establecimiento de reforzadores sociales ante la obtención de buenas calificaciones (r.0,287 con la mejora en comprensión; y r.0,298 con la mejora en APSL). Por último, hallamos la única correlación significativa (r.0,518) entre la mejora en comprensión y una respuesta que muestra la ausencia de reforzamientos: *aún no le he premiado*. Esta respuesta también mantiene correlación positiva con la escala de ajuste (r.0,258).

El segundo de los bloques de cuestiones contemplados en nuestro sistema de categorías, hace referencia a las ideas que tiene la familia acerca de la implicación en la actividad educativa escolar, así como en las posibilidades de influencia en la mejora de las dificultades de aprendizaje. Los análisis efectuados muestran ausencia de relaciones significativas entre las puntuaciones de mejora y las respuestas relativas a la implicación de la familia en la educación escolar. Ya pudimos comprobar, en este sentido, que la gran mayoría de las respuestas hacían referencia a la conveniencia de la colaboración y de la implicación familiar para contribuir a la mejora del aprendizaje. Este tipo de respuestas correlacionan positivamente con la mejora en la escala APSL (r.0,262); sin

embargo, las respuestas que aluden a la propia falta de capacidad para enseñarle, así como a la búsqueda de apoyo externo, correlacionan negativamente con la mejora en la escala de ajuste ($r=-0,515$).

Encontramos relación positiva, aunque no significativa ($r=0,507$) entre la mejora en APSL y las respuestas que hacen referencia a la influencia ejercida mediante el modelo/ejemplo: *si ven que en la casa hay bastantes libros y los padres leen, esto también ejerce influencia*. Encontramos igualmente correlación positiva entre la puntuación de mejora en APSL y las respuestas que hacen referencia a la influencia ejercida mediante ayuda o estimulación directa ($r=0,398$). Por su parte, algunas respuestas que hacen referencia a esta influencia ejercida mediante ayuda o estimulación directa, obtienen también correlación positiva con la mejora en comprensión, aunque sin llegar a la significatividad: *sí, animándole a leer lecturas que le gusten* ($r=0,196$); *sí, diciéndole que estudie, que intente leer de vez en cuando y aconsejándole que es bueno para expresarse* ($r=0,368$). Por otra parte, las respuestas que muestran ausencia de ideas claras en este sentido o que delegan esta influencia, *haciendo lo que diga la maestra*, obtienen valores de correlación negativa con las distintas puntuaciones de mejora, llegando a ser significativas con las de inteligencia ($r=-0,280$ con comprensión; $r=-0,315$ con APSL; y $r=-0,734$ con la prueba de CI). Igual resultado obtiene la relación entre algunas respuestas que sitúan la posible influencia en un ámbito material: *regalándole libros*, y las distintas mejoras ($r=-0,200$ con comprensión; $r=-0,315$ con APSL; y $r=-0,734$ con Cattell-2).

En cuanto al tercer bloque de cuestiones, aquel que aborda las expectativas y deseos relativos a los estudios o la profesión que desarrollarán sus hijos, encontramos una correlación positiva y significativa entre la mejora en comprensión y la expresión de expectativas positivas, aunque relacionadas con la situación de dificultades de aprendizaje, o que expresan de algún modo el conocimiento del deseo de los hijos: *lo que más le guste o lo más conveniente para él; es su opción* ($r=0,525$). Esta misma respuesta correlaciona positivamente con la mejora en la escala APSL ($r=0,215$). Encontramos igualmente correlación positiva entre otras respuestas incluidas en la categoría citada, y la mejora en APSL: *ella quiere ser pediatra o enfermera; me gustaría que lo fuera* ($r=0,507$). Por lo que respecta a las expectativas negativas, en general obtienen correlaciones negativas con las puntuaciones de mejora, aunque en algún caso en el que se desconocen los deseos de los hijos, hayamos relaciones positivas, aunque carentes igualmente de significatividad. En general, podríamos destacar que encontramos mayor número de correlaciones positivas, tanto con las puntuaciones de mejora obtenidas en las pruebas de comprensión como con las obtenidas en la escala APSL, en aquellas respuestas que hacen referencia a profesiones concretas o a deseos que hacen énfasis en la felicidad con aquello que el sujeto elija por sí mismo ($r=0,402$ con comprensión; y $r=0,507$ con APSL); que con respuestas que enfatizan aspectos de desarrollo personal-social, como ser una persona culta, educada, de provecho, etc. En cualquier caso, ninguna de las relaciones encontradas llega a ser significativa.

Con relación al cuarto bloque de cuestiones, que hace referencia a los elementos materiales relacionados con la lectura, no encontramos una relación significativa entre la disposición material de libros y las puntuaciones de mejora obtenidas. Hallamos una relación significativa ($r=0,568$) con una respuesta que indica sólo una pequeña cantidad: *cinco o menos de cinco*, mientras que hay una correlación negativa y también

significativa ($r=-0,531$) con la respuesta: *muchos*. Sin embargo, es destacable una relación positiva entre la mejora en APSL y algunas de las respuestas que aluden de forma concreta al tipo de libros infantiles o juveniles que se tienen: *libros de lectura; diccionarios; novelas* ($r=0,507$ en todos los casos).

En cuanto al quinto bloque de nuestro sistema de categorías, aquel que hace referencia a la experiencia de las familias con la lectura, lo más destacable es la ausencia de relaciones significativas entre aquellas respuestas que denotan un mayor contacto con la lectura, y la obtención de mejores resultados en las pruebas criterio.

El sexto bloque de cuestiones aborda las ideas sobre participación conjunta en actividades de lectura o de carácter audiovisual. El análisis de los resultados nos muestra algunas relaciones positivas entre respuestas favorecedoras de la participación y de la inclusión de los hijos en la actividad ($r=0,241$ con comprensión; y $r=0,215$ con APSL). Por lo que respecta a cómo se valora la participación conjunta en actividades de lectura como fuente de desarrollo cognitivo y lingüístico, hallamos relaciones positivas, aunque igualmente carentes de significatividad, entre las respuestas que muestran acuerdo con la idea expuesta en la cuestión, situando el énfasis en aspectos positivos ($r=0,317$ con la mejora en APSL; y $r=0,318$ con la mejora en CI). También se obtienen correlaciones positivas con respuestas que muestran acuerdo inicial, pero añaden alguna premisa o sitúan el énfasis en las dificultades: *es buena idea, aunque debe salir de ellos mismos* ($r=0,368$). Por otra parte, encontramos correlaciones negativas en algunas de estas respuestas que inciden en las dificultades de llevar a la práctica actividades de participación conjunta: *muy bien, pero es para tener mucho tiempo* ($r=-0,280$ con la mejora en comprensión; $r=-0,315$ con la escala de ajuste APSL; y $r=-0,734$ con la mejora en CI, valor éste que llega a ser estadísticamente muy significativo).

Con relación a los factores que influyen en las dificultades de aprendizaje, aspecto que se trata en el siguiente bloque de cuestiones, los resultados muestran una correlación positiva, aunque no significativa, entre la mejora en comprensión y en ajuste personal-social, y aquellas respuestas que destacan *la falta de atención* como uno de los factores fundamentales en la explicación de las dificultades en lectura ($r=0,493$ con comprensión; y $r=0,408$ con APSL). Hemos encontrado así mismo otros valores de correlación que pueden resultar de interés, aunque carezcan de significatividad estadística. En este sentido, otros aspectos que también centran las dificultades en el propio alumnado, como *la falta de esfuerzo*, correlacionan negativamente con la mejora en comprensión ($r=-0,448$), pero paradójicamente mantienen un nivel de correlación positiva con la escala de ajuste ($r=0,283$). En la misma línea de respuestas, encontramos una que hace referencia a procesos de autorregulación: *le cuesta concentrarse* (control de la atención), manteniendo correlación positiva con la mejora en la escala APSL ($r=0,398$). Por su parte, algunos aspectos que centran las dificultades de aprendizaje en el contexto familiar: *necesita más apoyo*, correlacionan positivamente con la mejora en comprensión ($r=0,368$). En cambio, otras respuestas relacionadas igualmente con el contexto familiar, pero que hacen referencia a situaciones de desventaja socio-cultural como factores explicativos de las dificultades, correlacionan negativamente con las distintas puntuaciones de mejora, llegando ser significativos los valores obtenidos con relación a la escala de ajuste ($r=-0,188$ con comprensión; $r=-0,611$ con la mejora en ajuste; y $r=-0,278$ con la mejora en CI).

Los resultados obtenidos con anterioridad vuelven a confirmarse en otra de las cuestiones de este mismo bloque, en la que encontramos una relación significativa entre las puntuaciones de mejora en comprensión ($r.0,595$) y en la escala de ajuste ($r.0,563$), por una parte, y las respuestas que hacen referencia a los factores relacionados con el propio alumnado, y más concretamente con *la falta de atención* como elemento fundamental en la explicación de las dificultades, por otra. La *falta de interés/motivación del propio alumnado*, como factor explicativo, también correlaciona positivamente con la mejora en ajuste, aunque no de forma significativa ($r.0,408$), así como con la mejora en la prueba de CI ($r.0,268$). Por otra parte, un aspecto que hace alusión al contexto escolar también correlaciona positivamente con la mejora en comprensión, aunque tampoco llega a ser significativo ($r.0,368$): *poco trabajo continuado en la escuela sobre lectura*. Sin embargo, otra respuesta similar, que también destaca el contexto escolar, y en concreto *factores relacionados con el profesorado y el sistema de enseñanza*, correlaciona negativamente con las distintas puntuaciones de mejora ($r.-0,280$ con comprensión; $r.-0,315$ con APSL; y $r.-0,734$ con la mejora en CI, valor que llega a ser estadísticamente muy significativo). Las respuestas que hacen referencia a factores relacionados con el contexto social, especialmente a *la competencia que suponen los medios audiovisuales*, obtienen valores de correlación negativa con la mejora en comprensión ($r.-0,448$), aunque estos valores pasan a ser positivos con la mejora en la escala de ajuste ($r.0,283$) y con la mejora en CI ($r.0,155$). Por último, nos parece de interés que algunas de las respuestas que apuntan como factores explicativos de las dificultades, aspectos relacionados con el contexto escolar, y más concretamente con el proceso de enseñanza-aprendizaje, correlacionen positivamente con la puntuación de mejora en la escala de ajuste personal-social: *los textos, temas y actividades relacionados con la lectura deben gustarles* ($r.0,507$).

El octavo bloque de cuestiones contiene 3 preguntas que abordan la concepción de la lectura y su relevancia. Encontramos relaciones positivas, aunque carentes de significatividad, entre las puntuaciones de mejora en comprensión y algunas respuestas que indican la validez de la lectura para el futuro y para desenvolverse en la vida: *para que en un futuro pueda expresarse correctamente* ($r.0,368$); *porque es imprescindible para la convivencia* ($r.0,368$); *leer es básico para desenvolverse en la vida* ($r.0,257$ con comprensión; y $r.0,326$ con la escala APSL). Otras respuestas que indican validez para el aprendizaje y para el enriquecimiento personal, también correlacionan positivamente con la mejora en APSL: *para aprender todo lo necesario en un futuro* ($r.0,308$). En general hallamos valores de correlación más altos con las respuestas que destacan la validez de la lectura para el futuro y para desenvolverse en la vida, que con las respuestas que apuntan a la validez de la lectura como aspecto fundamental para cualquier aprendizaje.

En cuanto al bloque de cuestiones que hace referencia a las expectativas sobre el aprendizaje de alumnos y alumnas que muestran dificultades de aprendizaje, los resultados muestran relaciones poco significativas entre las puntuaciones de mejora y las distintas respuestas. Entre éstas, las que expresan expectativas positivas sobre el aprendizaje de sujetos que comienzan su escolarización presentando dificultades en el aprendizaje de la lectura, haciendo referencia a que *con empeño pueden conseguir superar las dificultades*, correlacionan negativamente con la mejora en comprensión ($r.-0,448$), aunque muestra valores positivos de correlación con la mejora en APSL ($r.0,283$) y con la mejora en Cattell 2 ($r.0,155$). Así mismo es positiva la correlación con respuestas que expresan expectativas positivas, pero ponen el acento en que *le costará*

mucho (r.0,211). Por último, encontramos igualmente una relación positiva entre la mejora en APSL (r.0,507) y una respuesta que destaca la importancia conjunta del propio esfuerzo con el apoyo de los padres: *sí [podrá aprender a leer], si pone de su parte y los padres le apoyan.*

El último bloque de cuestiones valora la importancia otorgada a la colaboración con el profesorado de los hijos. Los resultados, al igual que ya ocurriera para el profesorado, muestran una relación positiva (r.0,368) entre la mejora en comprensión y la puntuación 3 otorgada por las familias; sin embargo, para la puntuación 4 encontramos valores de correlación negativos con la mejora en comprensión (r.-0,041) y en APSL (r.-0,647), siendo este último estadísticamente significativo. Por su parte, la puntuación 5 correlaciona negativamente con la mejora en comprensión (r.-0,469), pero obtiene un valor de correlación positivo con la escala de ajuste (r.0,390).

DISCUSIÓN Y CONCLUSIONES

10. Discusión

El presente capítulo tiene por objeto retomar, a la luz del marco teórico establecido, determinados aspectos especialmente relevantes en el desarrollo de nuestra investigación, algunos de los cuales ya han sido objeto de un cierto análisis en capítulos anteriores. Así mismo, nos proponemos introducir otros elementos; ya sea con relación a las bases teóricas y metodológicas del estudio, o relativos a los resultados obtenidos, con el fin de profundizar en el análisis de las razones que podrían ayudarnos a explicar mejor el proceso desarrollado y los cambios producidos. La presentación del contenido de esta sección pretende de alguna forma involucrar al investigador en la revisión y el análisis del proceso de evaluación seguido, o de los resultados obtenidos. Con este propósito, la organización del capítulo se ha establecido —en lo posible— teniendo en cuenta la estructura del informe de investigación que presentamos. De este modo, se propone en primer lugar la discusión general y el análisis de diversos aspectos que guardan relación con los aspectos teóricos y metodológicos de base; posteriormente se abordarán la discusión del proceso de aplicación experimental del dispositivo; el análisis sobre el propio proceso de evaluación de la aplicación y sobre los instrumentos de evaluación; la discusión de los resultados en las distintas pruebas que conformaban la batería criterio; y, finalmente, la discusión de los resultados de la evaluación complementaria.

10.1 *Discusión general relativa a las bases teóricas y metodológicas del dispositivo de evaluación dinámica EDPL*

Nos disponemos en primer lugar a revisar algunos aspectos relacionados con la *teoría de base* del dispositivo de evaluación dinámica EDPL. Esta revisión la efectuaremos en función de las propuestas teóricas y metodológicas que hemos analizado en el transcurso de la investigación, así como de los propios análisis sobre el proceso y los resultados obtenidos en la aplicación experimental. De este modo, iremos evaluando y contrastando las diversas opciones tomadas con relación a aquellos elementos que han sido considerados cruciales a lo largo del desarrollo de la investigación presentada. Estos elementos hacen referencia fundamentalmente a: (a) los objetivos de la evaluación dinámica, (b) la metodología empleada, (c) las estrategias de análisis de las pruebas dinámicas, (d) la medición del cambio, y (e) la validez ecológica de los resultados. Profundizaremos inicialmente en un enfoque evaluativo de las pruebas dinámicas, propuesto por Sternberg y Grigorenko (2003), para, posteriormente, ir situando y analizando nuestras decisiones y nuestro propio proceso de evaluación.

En su revisión de los modelos dinámicos de evaluación, estos autores observan que una posible explicación de la incoherencia que se observa en algunos de los resultados obtenidos en los análisis de los diferentes modelos, tendría que ver con la multiplicidad de objetivos que presiden algunos de éstos. Citando a Embretson (1987), hacen alusión a tres tipos de objetivos presentes en las propuestas dinámicas. El primero de ellos hace referencia a que la evaluación dinámica *debería aportar una mejor valoración del constructo de una aptitud determinada*. En este sentido, las pruebas dinámicas servirían

para incrementar la posibilidad de establecer comparaciones individuales más óptimas, eliminando o minimizando ciertos inconvenientes; un ejemplo de esto lo encontramos cuando un modelo dinámico busca igualar la base de partida de los sujetos que afrontan la prueba. La mayor igualdad inicial entre los sujetos de los contrastes podría incrementarse, según Sternberg (1997b) o Kozulin y Garb (2001), mediante la formación previa en aquellos contenidos o procesos que serán objeto de análisis, o bien proporcionando los resultados referentes a procesos previos necesarios para la resolución de los ítems de las pruebas; en este caso, la comparabilidad de los sujetos se estima mayor, y en los resultados obtenidos tendría un menor peso la historia individual. Como consecuencia del procedimiento dinámico de evaluación, y en sintonía con la idea expresada anteriormente, este objetivo de las pruebas dinámicas también hace referencia al incremento de la posibilidad de establecer mejores comparaciones individuales mediante la toma en consideración de las diferencias individuales de partida, es decir, teniendo en cuenta las diferencias culturales o de capacidad de los sujetos o los grupos en los análisis realizados. Igualmente, la consecución de este primer objetivo haría posible la reducción de los efectos de algunas variables relacionadas a menudo con el afrontamiento de pruebas psicológicas o de rendimiento, como la ansiedad. Sin embargo, para que la modificación de las condiciones antes mencionadas repercuta realmente en un incremento de la comparabilidad de los sujetos y, como consecuencia, de la capacidad de la prueba para ofrecer una mejor valoración del constructo de una aptitud, se requiere, por una parte, un análisis detallado de los efectos que producen en el análisis la modificación y la toma en consideración de los elementos o las diferencias mencionadas. Por otra parte, también sería necesario que las pruebas mostrasen un valor de predicción superior a las pruebas estáticas con relación a la aptitud estudiada, así como un conocimiento detallado sobre los efectos del modelo en factores relacionados con las aptitudes; desde aspectos relativos a las estrategias de aprendizaje, a otros más relacionados con los procesos de ajuste personal-social.

El segundo de los objetivos a los que Embretson hace referencia es la *medición de una nueva función psicológica*, como por ejemplo la modificabilidad cognitiva o el potencial de aprendizaje. En este caso, estas funciones son consideradas a modo de aptitudes independientes, y el objetivo de la evaluación dinámica se convierte en ofrecer un índice de estas nuevas aptitudes. Por último, el tercer objetivo de las pruebas dinámicas sería *la mejora de la eficacia cognitiva del sujeto*. En este caso, los procedimientos de evaluación dinámica son transferidos a procesos de instrucción con el objetivo de modificar y optimizar las funciones y la actuación cognitiva del sujeto. La evaluación dinámica se entremezcla de esta forma con la intervención.

Por otra parte, a nuestro entender, esta distinción de las pruebas dinámicas en función de los objetivos fundamentales que persiguen, parece supeditarse al *quién* y al *para qué* se utiliza un procedimiento de evaluación dinámica. En este sentido, el uso de modelos dinámicos de evaluación por parte de orientadores o profesores se tornará probablemente distinto al uso de pruebas dinámicas en el contexto de una investigación. Los primeros buscan identificar cuáles son y dónde se encuentran las dificultades de aprendizaje del alumno, al tiempo que tratan de determinar cómo pueden orientar y optimizar el aprendizaje; los investigadores buscan sobre todo analizar las causas del cambio cognitivo, el establecimiento de valores de predicción de las pruebas o la cuantificación del potencial de aprendizaje. Lógicamente, estos objetivos presiden y determinan en gran parte qué actividades se realizan y cómo.

En nuestro caso, la construcción del dispositivo EDPL tuvo lugar con el objetivo de ser útil al profesorado y a los equipos de orientación educativa que realizan primordialmente su trabajo en el contexto escolar con alumnado que presenta dificultades de aprendizaje en la lectura. Esto quiere decir que, en la práctica, sus objetivos van más allá de la cuantificación de un determinado indicador o del establecimiento de las causas del cambio producido. La evaluación propuesta pretende ser un instrumento eficaz para la mejora, es decir, está orientada a la solución de los problemas. En este sentido, estamos en sintonía con Swanson cuando expone entre sus preocupaciones el hecho de que el instrumento de evaluación construido pueda ser útil y eficaz para el trabajo que deben desarrollar los psicólogos educativos en los centros escolares (Swanson, 2000). En la misma línea, una revisión de los análisis cualitativos efectuados por Resing (2000) a raíz de la implementación de la prueba dinámica LIR, permitía extraer información sumamente valiosa desde una perspectiva centrada en la evaluación como proceso orientado a la mejora del proceso de enseñanza-aprendizaje. Ya apuntábamos con anterioridad que de poco serviría al profesorado preocupado por la mejora del funcionamiento cognitivo, el conocimiento aislado de un nuevo índice que viniera en la práctica a sustituir al CI (Saldaña, 2001). En este sentido, en los últimos años se ha observado un interesante cambio en la investigación sobre pruebas dinámicas, que apunta a la consideración, precisamente, de las características esenciales que tiene la aplicación de la evaluación dinámica en contextos como el educativo. El énfasis se ha puesto, en numerosos estudios, no tanto en la predicción o la cuantificación del potencial de aprendizaje, cuanto en la orientación a la intervención; no tanto en perfeccionar las pruebas de inteligencia, como en analizar el proceso de aprendizaje (Sternberg y Grigorenko, 2003). La consideración de estos factores de contexto y más relacionados con la aplicación práctica de las pruebas dinámicas, dificulta sin duda los procesos de investigación, pero, por otra parte, parece ofrecer datos más cualificados para la intervención y la mejora de los procesos educativos. Nuestro estudio ha tratado de alinearse en esta tendencia. Por otra parte, hacíamos referencia cuando analizábamos las bases teóricas y metodológicas de la propuesta de Jensen (1992, 1998, 2000), a que ésta no fue desarrollada con la intención de clasificar mejor al alumnado que presenta necesidades educativas especiales; ni, así mismo, para ofrecer una valoración del logro de un determinado sujeto con relación a un grupo. En este sentido, nuestro dispositivo de evaluación dinámica comparte básicamente estos puntos con el enfoque Mindladder. En efecto, EDPL no ha sido pensado ni construido con el objetivo de organizar o de clasificar mejor al alumnado que presenta dificultades; y tampoco fue desarrollado para situar a un sujeto con relación a un grupo, en función de su rendimiento en la prueba.

Otras conexiones que merece la pena destacar entre la propuesta de Jensen y la que aquí presentamos, hacen referencia a la consideración de los instrumentos desarrollados, como instrumentos de evaluación-intervención (Jensen, 2000). En efecto, al margen de que nuestra propuesta contemple los procesos de evaluación e intervención como procesos interdependientes que se retroalimentan de forma continua con el objetivo de optimizar el proceso de enseñanza y aprendizaje; el contenido conceptual y procedimental del dispositivo EDPL puede ser considerado también una base para el diseño y la implementación de actividades de aprendizaje contextualizadas en el ámbito escolar. Por otra parte, la teoría del *constructivismo mediado*, desarrollada por Jensen, propone la ED como parte integrante del proceso de enseñanza y aprendizaje en el aula, y perseguiría los mismos objetivos que la intervención, es decir, la optimización de las funciones de construcción del conocimiento. En este sentido, este autor hace referencia

a las dificultades existentes para establecer puentes entre los contenidos que suelen desarrollarse en las aplicaciones de pruebas de ED, y el currículum escolar. De esta forma, el aprendizaje efectuado por los sujetos en las sesiones de aplicación dinámica, no se mostraría relevante ni útil en otros contextos de aprendizaje. Los estudios desarrollados por este autor y sus colaboradores sugerían la necesidad de incluir el trabajo de estimulación cognitiva en el contexto de desarrollo y aplicación de estos procesos, es decir, en el contexto educativo. Jensen sugiere, por tanto, al igual que nosotros, la contextualización del proceso y de los instrumentos de evaluación. Como parte de esta contextualización en el marco escolar, la propuesta de este autor contempla la colaboración directa del profesorado de los sujetos que son objeto de valoración, incluyendo esta colaboración algunos de los aspectos que también han sido considerados en nuestra propuesta de evaluación dinámica. Estos aspectos son relativos al intercambio de información, a la valoración conjunta de las dificultades del sujeto y de los resultados del proceso de evaluación, así como a la determinación de orientaciones psicopedagógicas contextualizadas con base en el proceso de evaluación dinámica efectuado.

La aplicación experimental del dispositivo dinámico EDPL habría conectado evaluación e intervención. Las diversas actividades propuestas, así como la puesta en práctica de los patrones metodológicos de evaluación dinámica utilizados en la mediación, habrían contribuido al ajuste continuo de la intervención del aplicador a las características y necesidades del sujeto evaluado, situándose de esta forma en su *zona de desarrollo potencial*. La aplicación del dispositivo EDPL ha evaluado, entre otros aspectos: cómo se produce el aprendizaje, mediante qué mecanismos tiene lugar este aprendizaje, y cuál es la calidad del mismo y de los cambios producidos; todo esto con el objetivo primordial de optimizar el proceso de aprendizaje, al tiempo que orientar hacia una intervención educativa que continúe y mejore el desarrollo del potencial de aprendizaje. Así mismo, habría obtenido información sobre aquellos procesos cognitivos, metacognitivos y de ajuste personal-social, que los sujetos experimentales han utilizado en la resolución de las tareas propuestas, determinando las condiciones que optimizarían su puesta en práctica. El proceso de evaluación dinámica aplicado se ha desarrollado en el contexto escolar, es decir, en el contexto donde tiene lugar la acción educativa encaminada a la optimización de los procesos de lectura y comprensión. Ha obtenido información fundamentalmente sobre el proceso de enseñanza y de aprendizaje relacionado con la lectura, y lo ha hecho en situaciones marcadas por la interacción y la colaboración. Por otra parte, podemos decir que el procedimiento de evaluación-intervención puesto en práctica, contiene actividades orientadas a la obtención de información valiosa sobre el funcionamiento cognitivo; pero también a la *mejora* de la actuación cognitiva. En este sentido, dado que la distancia entre las actividades propuestas y el contenido de las tareas criterio pretest y postest, es considerable, las ganancias obtenidas por los sujetos pueden ser consideradas como muy relevantes; estas ganancias no expresarían sólo diferencias pretest-postest, sino que también serían indicativas de la existencia de procesos de transferencia «lejana» (Campione y Brown, 1987; Campione, 1989). Los resultados indicativos de mejora obtenidos en el postest, otorgan ciertamente información sobre la incorporación o integración de la información, y/o sobre los procesos entrenados que han debido ser puestos en práctica en la ejecución de las actividades. Adicionalmente, los resultados obtenidos del análisis del proceso de aplicación, especialmente del análisis de los patrones metodológicos de evaluación dinámica, nos otorgan información potencial sobre la propensión y la capacidad de un

sujeto para beneficiarse de la mediación ejercida por los aplicadores con el objetivo de mejorar su actuación; es decir, nos informa de su potencial dinámico de aprendizaje.

Por otra parte, la coherencia con las ideas sobre evaluación e intervención defendidas en nuestro trabajo, implicaba la necesaria participación activa y mediadora de los profesionales que realizaron la aplicación; sólo de esta forma hemos podido obtener la información relativa a los elementos metodológicos o contextuales que podrían haber contribuido a optimizar el rendimiento de los sujetos, facilitando su aprendizaje. Igualmente, sólo desde la puesta en práctica de los patrones metodológicos de evaluación dinámica, utilizados en el proceso de mediación, hemos podido acceder a información valiosa sobre procesos y estrategias que, de otra forma, podrían haber quedado solapados en valoraciones de corte más estático. Se ha tratado, tal y como recogen Sternberg y Grigorenko (2003), de revelar el potencial oculto del sujeto en el transcurso del proceso de evaluación.

Hemos tratado, así mismo, de ser fieles en la medida de lo posible a aquellos aspectos considerados significativos y determinantes para la obtención de información relevante acerca de los procesos estudiados, así como para poder realizar propuestas de mejora que se muestren sensibles al contexto educativo. Aunque en nuestra propuesta original, recogida en buena parte a lo largo de estas páginas, se explicita de forma clara la conveniencia, y el modo, de aproximarse a la evaluación del contexto educativo donde se desarrolla el proceso que evaluamos, en nuestra aplicación experimental del dispositivo EDPL no hemos podido cumplir con todas las premisas expuestas con anterioridad. Ya comentamos la imposibilidad por parte de la totalidad de los aplicadores, de realizar la actividad 32 del dispositivo, que hubiera otorgado la oportunidad de analizar de una forma más sistemática y profunda el contexto de aula en cada una de las situaciones estudiadas. Sin embargo, a lo largo del desarrollo de la aplicación, hemos tenido oportunidad de recabar abundante información relativa al contexto; tanto al contexto de aula, como al socio-familiar del alumnado. En este sentido, hemos podido acceder al conocimiento de circunstancias o condiciones de carácter socio-cultural a través de los profesores-colaboradores, incluyendo estas circunstancias entre las variables tenidas en cuenta a la hora de efectuar el análisis de los resultados. Igualmente, hemos podido aproximarnos al conocimiento sobre la realidad cotidiana, tanto escolar como socio-familiar, en la que ha tenido lugar el proceso de enseñanza y aprendizaje, así como el proceso de aplicación experimental del dispositivo, a través del análisis de los cuestionarios EIDAL y EIDAL-MP.

En línea con algunos de los argumentos expuestos en las últimas páginas, relativos a la contextualización del proceso y de los instrumentos de evaluación, compartimos con Lidz (2000) la idea de que las pruebas de ED construidas con base en los programas escolares —como es nuestro caso—, deben trascender la valoración de la calidad de la consecución de unos objetivos educativos específicos relacionados con las áreas escolares. Deberían, en cambio, poder ofrecer información adicional, sobre todo con relación a la capacidad de aprendizaje del alumno evaluado en tareas relacionadas con los contenidos escolares. Así mismo, también deberían informar sobre las dificultades que el sujeto muestra en el proceso de aprendizaje. De este modo, un proceso idóneo de evaluación debería contar con el desarrollo de pruebas de ED relacionadas con el currículum escolar y con los procesos que subyacen al aprendizaje de sus contenidos. Es en este sentido en el que nuestra propuesta de evaluación dinámica de procesos implicados en la lectura, comparte básicamente planteamientos y métodos de análisis

con la propuesta de Lidz, contemplando especialmente los aspectos mencionados. Por otra parte, tal y como recogíamos con relación a la prueba dinámica ACFS, de Lidz y Jepsen (1997), los resultados obtenidos con la aplicación de nuestro dispositivo, requerirían de estudios posteriores con el objetivo de valorar en qué medida nuestra propuesta de ED de procesos implicados en la lectura se muestra capaz de obtener una mayor validez predictiva sobre el rendimiento escolar, con relación a otras pruebas dinámicas de carácter descontextualizado, y, por tanto, ajenas a los programas y contenidos escolares. Estos estudios deberán valorar las predicciones realizadas sobre el rendimiento de los sujetos en actividades de lectura y comprensión, así como la incidencia de las orientaciones psicopedagógicas que se derivan de la aplicación dinámica de EDPL, y que son ofrecidas al profesorado.

Hemos repasado en estas últimas páginas algunas de las premisas fundamentales de nuestra propuesta de evaluación dinámica, relacionadas con el *por qué* y con el *cómo* de la misma; y centradas así mismo en una práctica contextualizada en el ámbito escolar. Sin embargo, en el contexto de nuestra investigación, nuestro interés ha debido encaminarse principalmente hacia el análisis. Ya hicimos mención de que algunas de las premisas fundamentales de nuestra propuesta inicial, fueron de alguna forma reducidas con el objetivo de incrementar el control experimental del estudio. En este sentido, en la introducción del capítulo de Instrumentos, se recogían algunas de estas consideraciones que hemos contemplado como *claves* en la conceptualización y el desarrollo de procedimientos dinámicos de evaluación, y que, de alguna forma, se han visto afectadas en nuestra propia investigación. En este sentido, cuando aludíamos a la fase de construcción del dispositivo y a su aplicación piloto, apuntábamos que, aunque éste no contenía un tiempo prefijado de aplicación, considerábamos que, en las condiciones habituales en las que se desarrolla una evaluación psicopedagógica en el contexto escolar, podría ser realizado en pocas sesiones, en función de la formación del evaluador y de la información previa con la que se contara sobre el alumno evaluado. En cambio, este análisis de la aplicación del dispositivo EDPL, tal y como ha sido concebido para su uso en el contexto escolar, es decir, primando aspectos como: la contextualización de los instrumentos de evaluación y la flexibilidad del evaluador a la hora de elegir: (a) qué actividades del esquema propuesto son las que mejor se adecuan al problema tratado, (b) qué recorrido realizará por ese esquema para lograr detectar dónde se encuentran las dificultades, y (c) cuál es el modo óptimo de mejorar los procesos afectados; planteaba algunos inconvenientes teóricos desde la óptica de la investigación experimental y académica. En efecto, la lógica habitual de la investigación académica, quizás todavía instalada en demasía en parámetros psicométricos, aconsejaba plantear un diseño en el que la aplicación experimental de nuestro dispositivo de ED contara con una serie de características que dificultaban seriamente la implementación del dispositivo, tal como había sido concebido. Entre estas características estaban las siguientes: (a) por una parte, parecía aconsejable que la duración de la fase experimental no fuera inferior a unos seis meses, con el objetivo de poder establecer comparaciones y posibles diferencias con la suficiente fiabilidad y validez entre las fases pretest y posttest; (b) así mismo, la aplicación del dispositivo debería realizarse de forma lo más homogénea posible. Como hemos podido examinar cuando hemos analizado las bases teóricas y metodológicas de nuestra propuesta, en el dispositivo de evaluación dinámica con el que han trabajado los profesores-colaboradores, se proponía la contextualización de los instrumentos de evaluación como la opción ideal; sin embargo, durante la fase experimental del estudio, convenía —a efectos de poder establecer comparaciones válidas entre las distintas aplicaciones

efectuadas, y controlar, en la medida de lo posible, variables relacionadas con aplicaciones excesivamente diferentes— efectuar una aplicación del dispositivo que se ajustara lo más posible a las pautas metodológicas recogidas en el manual que fue entregado a los aplicadores en las sesiones iniciales del Grupo de Trabajo, y siguiendo una secuencia prefijada con relación a las actividades propuestas.

Esta «contradicción», como la definíamos cuando analizábamos uno de los modelos dinámicos de evaluación, no ha dejado de plantearnos una cierta incomodidad y, a efectos prácticos, un cierto problema, ya que, en definitiva y a pesar de las justificaciones de orden teórico o práctico que hemos podido hallar, la valoración que hemos efectuado de la aplicación dinámica del dispositivo ha sido realizada en buena parte con pruebas estáticas de evaluación. Como hemos recordado, hacíamos referencia a este problema cuando abordábamos el análisis crítico de algunos de los modelos vigentes de aplicación dinámica. Apuntábamos entonces que el uso de pruebas de corte convencional y estático para valorar un proceso de evaluación dinámica podría conllevar una pérdida notable de información, debido fundamentalmente a que no podría obtenerse información sobre aspectos como: el *proceso* mismo de aplicación del dispositivo; sobre las dificultades encontradas en su puesta en práctica; los procesos de interacción mediada y los patrones metodológicos de evaluación dinámica que pueden haber contribuido a las mejoras observadas; los cambios y las rectificaciones en la respuesta de los sujetos ante los cambios en la presentación o formulación de las cuestiones; la transformación progresiva de las funciones cognitivas o socio-personales trabajadas; etc. En definitiva, sobre aquellos aspectos que pueden marcar y significar diferencias y mejoras significativas sobre los procedimientos de evaluación convencionales, contribuyendo a ampliar y enriquecer la información obtenida y facilitando la conexión del proceso de evaluación con la posterior acción educativa encaminada a optimizar el proceso.

Hemos tratado, sin embargo, de minimizar en lo posible esta pérdida significativa de información mediante la utilización de vías complementarias de valoración cualitativa, al tiempo que centradas en el proceso mismo de aplicación del dispositivo. Se ha destacado en no pocas ocasiones (Fernández-Ballesteros, 1992; Mora, 1998) que el empleo paralelo y complementario de distintos sistemas de recogida de información mediante actividades o pruebas de evaluación con características y naturaleza distinta, puede erigirse en una fortaleza de la investigación realizada, por cuanto aporta diversas visiones y perspectivas desde las que abordar el problema, compensa ciertos sesgos que de forma inevitable puede acarrear el empleo de pruebas de carácter cuantitativo o cualitativo exclusivamente, así como permite validar con una mayor potencia los resultados convergentes. De esta forma, a lo largo del desarrollo de la fase experimental del estudio, se han utilizado los siguientes sistemas de recogida de información complementaria: (a) Cada uno de los profesores-colaboradores ha recogido en los cuadernillos donde se encontraban las *Hojas de registro y evaluación* de cada una de las actividades propuestas, varios aspectos relativos a la valoración del proceso de aplicación: mediación efectuada; valoración de la puesta en práctica de los distintos procesos implicados, especialmente de los metacognitivos; dificultades encontradas para la aplicación; evaluación del proceso de resolución de la actividad por parte del alumno; etc. (b) Así mismo, estos profesores han cumplimentado un cuestionario de preguntas abiertas en el que se pretendía específicamente analizar el dispositivo EDPL y el proceso de aplicación experimental. (c) Se ha recogido información por escrito de cada sesión del Grupo de Trabajo donde participaron los profesores-colaboradores del

estudio; el contenido de estas sesiones ha sido primordialmente el análisis conjunto de la puesta en práctica de EDPL, con atención especial a la valoración del proceso de aplicación, destacando aquellos patrones metodológicos de evaluación dinámica que contribuirían a mejorar el proceso, así como a la valoración de la puesta en práctica de procesos metacognitivos por parte de los sujetos. (d) En estas mismas sesiones del Grupo de Trabajo a las que hacíamos referencia, se han analizado conjuntamente varias sesiones de video que habían sido grabadas por uno de los participantes. (e) Se ha recogido información sobre las ideas y expectativas del profesorado (tutores del alumnado participante y también profesores-colaboradores) y de las familias, acerca de las dificultades de aprendizaje de la lectura, a través de los cuestionarios EIDAL y EIDAL-MP.

Por otra parte, este enfoque de los modelos dinámicos, cercano a los procesos de intervención y mejora del aprendizaje, conlleva en cambio que su aplicación práctica se realice en un contexto con características que pueden dificultar los procesos de control y análisis. En efecto, en el contexto escolar, los educadores especializados deben trabajar a menudo con pequeños grupos de sujetos con dificultades, al tiempo que deben interpretar y responder de forma flexible e inmediata a las respuestas y reacciones del alumnado ante las tareas propuestas (Sternberg y Grigorenko, 2003). Estas tareas, por otra parte, difícilmente pueden ser disgregadas y analizadas previamente en una secuencia de pasos sucesivos e invariables, que siga a su vez un único camino trazado de antemano hasta llegar al logro del objetivo previsto. Sin duda, estas características dificultan la aplicación de modelos de evaluación dinámica, ya sea en el contexto de una investigación o en la aplicación práctica; sin embargo, pensamos que la investigación debe tomar en consideración este contexto real de actuación para poder ofrecer a su vez información con una mayor validez ecológica.

Otro aspecto relacionado con la validez del proceso y de los resultados obtenidos, tiene que ver con la elaboración de pruebas dinámicas más específicas y dirigidas a ámbitos particulares del aprendizaje (Kozulin y Garb, 2001). En efecto, en los últimos años se han desarrollado algunas tareas dinámicas específicas sobre aspectos como, por ejemplo, la conciencia fonológica, la memoria de trabajo o la comprensión lectora en una segunda lengua. Los resultados parecen mostrar que estas tareas tienen un valor predictivo superior a pruebas dinámicas del potencial de aprendizaje que evalúan aspectos generales, así como también a pruebas estáticas que evalúan la inteligencia (Hamers, Pennings y Guthke, 1994; Grigorenko, Sternberg y Ehrman, 2000). Así, la inclusión en las pruebas dinámicas de actividades propias del contexto escolar, podría facilitar una mayor conexión entre los resultados de la evaluación dinámica y los resultados del aprendizaje escolar. En este sentido, hemos desarrollado en nuestra propuesta actividades de ED para diversos ámbitos específicos del aprendizaje de la lectura, entre ellos la conciencia fonológica, el establecimiento de relaciones entre textos o la lectura de pseudopalabras. En realidad, todas las actividades diseñadas en el dispositivo están basadas en procesos de lectura y comprensión y, por tanto, en ámbitos específicos del aprendizaje. Consideramos que la contemplación de estos ámbitos específicos y relacionados con los aprendizajes escolares, en el diseño de las actividades del dispositivo EDPL, es necesaria para incrementar la validez ecológica.

Al hilo de lo expuesto anteriormente, entre los objetivos de la aplicación del dispositivo EDPL no ha figurado sólo la predicción de la actuación de los alumnos en el contexto escolar, sino, fundamentalmente, la evaluación orientada al tratamiento y la mejora de

los procesos cognitivos y socio-afectivos implicados en la comprensión de textos. También hemos pretendido contribuir a la posible introducción de cambios en los sistemas educativos relacionados con la evaluación y la intervención sobre los procesos de lectura. En este sentido, la constatación por parte del profesorado colaborador de las limitaciones de los procedimientos estáticos de evaluación para dar cuenta de la complejidad de los procesos de aprendizaje y dominio de la lectura, y sobre todo para orientar de forma adecuada la intervención educativa; unida a la evidencia de resultados positivos en la aplicación del dispositivo de ED empleado, podrían facilitar la utilización de procesos dinámicos de evaluación como alternativa a los métodos habitualmente empleados con los sujetos que presentan dificultades en su aprendizaje.

En síntesis, hemos presentado en nuestra investigación un análisis de la *teoría de base* de nuestra propuesta de evaluación dinámica de procesos implicados en la lectura, profundizando en las bases teóricas y metodológicas para la evaluación dinámica de estos procesos en el contexto escolar. A pesar de las limitaciones que han sido apuntadas en los últimos párrafos; las características fundamentales de nuestro dispositivo, relativas a la relevancia de la contextualización del proceso y de los instrumentos de evaluación, así como a la flexibilidad del proceso de aplicación; han caracterizado en lo posible, el diseño de investigación, la fase experimental de aplicación del dispositivo, así como el proceso de evaluación seguido. Hemos abordado en este análisis cuestiones teóricas y prácticas relacionadas con la posibilidad y las dificultades de establecer procedimientos dinámicos de evaluación en el contexto educativo, que reporten información valiosa y adicional a la que podría obtenerse con la aplicación de pruebas convencionales; así como cuestiones psicométricas relacionadas con las posibilidades de optimizar el análisis del cambio producido en los sujetos como fruto de su participación en procesos de ED. Hemos abogado, de acuerdo con Schlee (1985), por la relevancia de la validez ecológica como parámetro fundamental en la evaluación de la aplicación de procesos de ED en el ámbito escolar. Hemos centrado así mismo nuestros objetivos, más allá de la predicción, en la evaluación orientada a la mejora del proceso de enseñanza y aprendizaje, mediante la continua optimización del desarrollo potencial dinámico del aprendizaje. Hemos considerado de esta forma la ED como un proceso de evaluación del potencial dinámico del aprendizaje de un sujeto con el objetivo principal de orientar la acción educativa hacia la optimización del proceso educativo. Por último, hemos sintetizado y recogido una serie de elementos que consideramos fundamentales para optimizar el proceso de evaluación; estos elementos a tener en cuenta son: la evaluación del contexto de aula, la contextualización de los instrumentos de evaluación, la evaluación dinámica preferente de procesos metacognitivos y la evaluación del proceso de resolución de las tareas planteadas.

10.2 *Discusión sobre el proceso de aplicación experimental de EDPL*

Estructura, contenidos e información analizada en el dispositivo EDPL

En nuestro análisis de la aplicación experimental de EDPL, se ha evaluado la *estructura* del dispositivo, así como los *contenidos* del mismo. En este sentido, hemos hecho alusión a los procesos implicados en la comprensión de textos que han sido contemplados en el dispositivo, así como a los indicadores que finalmente han servido al profesorado aplicador como criterios de evaluación durante las sesiones de aplicación de EDPL. Se ha analizado también el registro de los elementos relevantes en la evaluación del proceso de resolución de las actividades propuestas y de los resultados.

Estos elementos hacían especialmente alusión al proceso de resolución de las actividades, al análisis de las posibles dificultades encontradas, a la mediación efectuada por los aplicadores, y a la puesta en práctica de estrategias metacognitivas por parte del alumnado. El análisis efectuado sobre el contenido de las *Hojas de registro y evaluación*, tuvo como ejes fundamentales el *tipo de información recogida y destacada por los aplicadores* y los *patrones metodológicos de evaluación dinámica utilizados en la mediación*.

En cuanto al tipo de información que los aplicadores destacaron en sus observaciones, identificamos finalmente 24 unidades de información significativas. Con este análisis queríamos conocer qué aspectos habían sido considerados fundamentales por los aplicadores, y si éstos guardaban relación con los resultados obtenidos en cada uno de los grupos de aplicación. Los análisis realizados nos permitieron comprobar que los aplicadores centraron su atención y sus observaciones en los procesos considerados *clave* en nuestra propuesta de evaluación dinámica, es decir: los procesos metacognitivos y el proceso mismo de mediación. En este sentido, el análisis de las correlaciones efectuadas nos mostró que las anotaciones de los aplicadores relativas al proceso de mediación, hicieron referencia de forma mayoritaria a la *necesidad de mediación* que mostraron los sujetos. Por otra parte, también pudimos comprobar que el número de anotaciones efectuadas sobre la puesta en práctica de un determinado proceso, no guardaba de forma general una relación positiva y significativa con la obtención de puntuaciones dinámicas más altas, lo que venía a indicar que el análisis y la valoración de los indicadores propuestos, efectuada por los aplicadores, contempló la calidad de la puesta en práctica de los procesos y no solo la presencia del indicador. Esto se vería reflejado igualmente en los niveles de correlación positiva alcanzados entre la puesta en práctica de la mayoría de los procesos contemplados y la *necesidad de mediación*. De esta forma, las anotaciones realizadas sobre la puesta en práctica de los procesos, habrían reflejado mayoritariamente la presencia de dificultades y la necesidad de la mediación del profesor, dato que quedaba confirmado al certificar la correlación negativa y significativa entre la necesidad de mediación y las puntuaciones dinámicas obtenidas.

Otro dato que resulta de nuestro interés es que el valor de correlación obtenido entre la *mejora durante el proceso de mediación* y la puntuación dinámica global, puede resultar contradictorio en primera instancia. En efecto, en buena lógica, cabría esperar que las anotaciones de los aplicadores relativas a la mejora del sujeto durante el proceso de mediación, se correspondieran con puntuaciones dinámicas más altas. Sin embargo, a pesar de que, efectivamente, las anotaciones del aplicador reflejaron la mejora de los sujetos durante el proceso, esto no se tradujo necesariamente en buenas puntuaciones globales del proceso. Es posible que algunos sujetos hayan experimentado mejoras, más o menos puntuales, durante el desarrollo de las actividades, y que éstas no se hayan visto reflejadas finalmente, o lo hayan hecho con menos intensidad, en la valoración realizada sobre los indicadores de evaluación. Dicho de otra forma, aunque un sujeto experimentara mejoras en el transcurso del proceso de mediación, pudo igualmente haber puntuado bajo en función de la valoración sobre la puesta en práctica de los indicadores propuestos. Esta hipótesis puede comprobarse, ya que las anotaciones sobre la mejora durante el proceso, han estado ligadas a la mayor necesidad de mediación de los sujetos. Como podemos ver, la correlación de la mejora durante el proceso, con la necesidad de mediación, es positiva y muy significativa, con lo que los sujetos que más apoyo necesitaron —que son los que obtuvieron finalmente puntuaciones más bajas—

fueron también aquellos sobre los que los aplicadores reseñaron mejoras durante el proceso de mediación. Por otra parte, los sujetos que necesitaron menos apoyo en la resolución de las tareas y puntuaron más alto, habrían obtenido, lógicamente, una menor cantidad de anotaciones relativas a mejoras durante el proceso.

Así mismo, también hemos analizado el conjunto de patrones metodológicos de evaluación dinámica utilizados en la mediación. Hemos efectuado un análisis minucioso sobre estos patrones metodológicos puestos en práctica por el profesorado, y hemos tratado de relacionarlos con los resultados obtenidos por los sujetos. Se han determinado, de esta forma, una serie de pautas docentes —en el sentido en el que Mora utiliza el término para hacer referencia a los *patrones de movilización cognitiva* en el marco del programa CyT (Mora, 1998; Aguilera y Mora, 2003)— que podrían haber resultado *claves* en las mejoras observadas. En este sentido, como recordaremos, el análisis efectuado por Sternberg y Grigorenko (2003) sobre el método de las sugerencias graduadas, desarrollado por Campione y Brown (Campione y Brown, 1987; Campione, 1989), contenía una crítica relativa a la posibilidad de que las distintas sugerencias pudieran tener diferentes efectos y significados en función del perfil cognitivo de los sujetos, confundiendo la medida del potencial de aprendizaje que se pretendía obtener, con las características relativas a la capacidad de atención, memoria o reflexión sobre las tareas propuestas. En este sentido, los patrones metodológicos de ED propuestos en EDPL deberían entrar en relación con estas características mencionadas anteriormente. De otro modo, las respuestas a la intervención, es decir, a la mediación efectuada, podrían reflejar aspectos de este tipo más que el aprendizaje en sí. Por otra parte, como ya destacábamos cuando abordábamos el análisis y la revisión crítica del modelo dinámico propuesto por Campione y Brown, puede no ser posible llegar a separar totalmente los distintos procesos cognitivos, metacognitivos y de ajuste personal-social implicados en una actividad como la lectura, del perfil cognitivo y socio-afectivo de los sujetos. De este modo, con relación al estudio presentado, consideramos, como ya apuntábamos entonces, que los efectos de estas variables relativas al sujeto, podrían haber quedado atenuadas en su mayor parte en función de las características de la muestra seleccionada, así como de los diversos análisis efectuados. En efecto, la muestra estaba compuesta en su mayoría por sujetos que compartían características como: presentar dificultades de aprendizaje de la lectura, desventaja socio-cultural, bajo rendimiento cognitivo, presencia de discapacidad o recibir apoyo especializado. Por otra parte, entre los distintos análisis efectuados, se han tenido en cuenta las posibles diferencias que habrían tenido lugar en los resultados como consecuencia de la inclusión de los sujetos en los varios subgrupos conformados. Igualmente, con relación al análisis cualitativo y complementario sobre el proceso de aplicación del dispositivo y sobre los resultados obtenidos, hemos tenido en cuenta, y analizado, distintos bloques de procesos cognitivos y metacognitivos; entre ellos, los procesos psicológicos subyacentes, que incluían aspectos relativos a la atención, la memoria o los procesos perceptivo-visuales.

Los análisis efectuados nos mostraron que la mayoría de los patrones metodológicos de evaluación dinámica correlacionaban negativamente con las diversas puntuaciones dinámicas consideradas en el dispositivo. La explicación de este hecho parece residir en que una necesidad mayor de mediación por parte de los sujetos, habría supuesto la presencia de mayores dificultades en los procesos evaluados. Estos datos contrastan con la evidencia de mejora y modificación en sentido positivo, que experimentaron algunos de los sujetos que necesitaron mediación y no obtuvieron altas puntuaciones. En efecto,

los valores de correlación obtenidos para las puntuaciones de mejora, revelaron la presencia de ganancias postratamiento en sujetos que necesitaron de la mediación de los aplicadores para la realización de las tareas. Estos resultados fueron especialmente importantes con relación a la prueba de comprensión de primaria, para la que pudimos observar un mayor número de valores de correlación positivos entre patrones de mediación y puntuación de mejora. Los valores de predicción de la puntuación de mejora para esta prueba ya nos habían indicado la relevancia de la puntuación dinámica en las ganancias observadas, lo que vendría a verificar lo crucial de la mediación ejercida por los aplicadores en la consecución de estas ganancias. Sin embargo, con relación a los sujetos de secundaria y a la prueba de comprensión, los datos nos mostraron que las mejoras observadas no guardaban relación aparente con el proceso de mediación, a tenor de las anotaciones efectuadas por los aplicadores en las *Hojas de registro y evaluación*.

Por otra parte, la revisión de los datos nos mostraba que el número de correlaciones negativas entre patrones de mediación y puntuaciones dinámicas, disminuía en función del grado de dificultad que los distintos procesos contemplados representaron para los sujetos. Como podemos recordar, las puntuaciones dinámicas de los distintos procesos contemplados mostraban que los procesos de asociación G-F y los procesos psicológicos subyacentes, eran los que habían representado una menor dificultad para los sujetos experimentales. Es lógico pensar, por lo tanto, que la mayor necesidad de ayuda o mediación en este tipo de habilidades haya supuesto la presencia de problemas de lectura más acusados. A medida que las habilidades requeridas representaban una mayor dificultad, la necesidad de mediación no se habría relacionado tan estrechamente con mayores problemas en la resolución de las tareas. La confirmación de estas primeras observaciones nos llega cuando analizamos los resultados de los análisis correlacionales para las puntuaciones dinámicas obtenidas en los procesos de integración texto-conocimientos previos. Estos procesos eran, con cierta diferencia, los que obtuvieron unas puntuaciones más bajas, siendo, por tanto, los que representaron una mayor dificultad para el alumnado participante. Podemos comprobar que la práctica totalidad de los valores de correlación mostrados por los distintos patrones de mediación, con la puntuación dinámica de estos procesos, son positivos y algunos de ellos llegan a la significatividad estadística. Este comportamiento diferencial, tal y como ya ocurriera en el análisis efectuado sobre el tipo de información recogida y destacada por los aplicadores, se relacionaría, como hemos apuntado, con las mayores dificultades que representó la puesta en práctica de los indicadores de evaluación de estos procesos. En este sentido, los sujetos «habrían necesitado» de la mediación del aplicador para acceder a una mejor ejecución de las actividades y, por tanto, a mejores valoraciones. Esta mediación habría sido en este caso necesaria y oportuna para todos los sujetos y no sólo para aquellos que más problemas manifestaban, obteniendo mejores puntuaciones quienes más y mejor se beneficiaron de ella.

Mientras que para las pruebas de comprensión y ajuste, encontramos un alto número de correlaciones negativas o de bajo nivel entre los patrones de mediación y las puntuaciones dinámicas, dato que, como hemos tenido oportunidad de ver, se relativiza cuando tomamos como referencia la puntuación de mejora; para la prueba de inteligencia, los datos nos mostraron algo bien distinto. En este caso, todos los patrones metodológicos de evaluación dinámica utilizados en la mediación correlacionaron de forma positiva con las mejoras observadas. Estos resultados revelarían la importancia

que el proceso de mediación efectuado habría tenido en la explicación de las ganancias observadas en esta prueba.

Por otra parte, los análisis efectuados nos llevaron a comprobar el comportamiento de los datos cuando eran seleccionados aquellos sujetos que obtuvieron mejoras postratamiento, es decir, el grupo de *mejoradores*. Ya expusimos cuando abordábamos la presentación de los resultados, que se produjo un incremento significativo en los valores de correlación observados entre los patrones de mediación y las puntuaciones dinámicas, especialmente en lo referente a los procesos de ajuste personal-social y a los procesos metacognitivos. Este incremento en los niveles de correlación tuvo lugar con relación a la prueba de comprensión ECO 1 y, muy especialmente, en la prueba de comprensión ECOS, para la que, además de producirse un incremento notable con respecto a los procesos citados, llegamos a encontrar valores de correlación positivos y muy significativos entre algunos de los patrones de mediación y los procesos de integración texto-conocimientos previos. Por lo que respecta a la escala de ajuste, pudimos comprobar igualmente que los niveles de correlación entre patrones de mediación y puntuaciones dinámicas, experimentaban un incremento significativo. Al igual que ocurriera en la prueba de comprensión para secundaria, este incremento tuvo lugar especialmente con relación a los procesos de integración texto-conocimientos previos, para los que también se hallaron niveles de correlación significativos con relación a algunos de los patrones de mediación utilizados. Por último, en referencia a la prueba de CI, se observaron igualmente incrementos muy notables en los valores de correlación que los patrones de mediación establecían con los puntajes dinámicos. En este sentido, algunos de los valores, ya positivos, pasaron a ser estadísticamente significativos. En todo este proceso, se revelaron algunos de los patrones de evaluación dinámica que habrían sido claves en la configuración de las mejoras observadas en las distintas pruebas; estos serían:

Para la prueba de comprensión en primaria, y tomando como referencia a la totalidad del grupo experimental: la *recapitulación conjunta del trabajo realizado*; la *mediación de la discusión y reflexión sobre las estrategias de comprensión*; la *mediación de la conciencia sobre las propias estrategias de comprensión*; y la *mediación del uso de estrategias metacognitivas de planificación y autoevaluación*. Cuando se tomaron como referencia a los sujetos mejoradores, los patrones de mediación relativos a los procesos metacognitivos, así como a los procesos de ajuste personal-social, como apuntábamos anteriormente, fueron los que, incrementando notablemente sus valores de correlación, se revelaron cruciales en la explicación de las ganancias observadas.

Para la prueba de comprensión en secundaria, tomando como referencia a la totalidad del grupo experimental, los patrones de mediación más importantes fueron: la *mediación de la conciencia de la estructura fonológica, léxica y sintáctica*; y la *mediación de la conciencia de las propias dificultades*. Cuando se tomó como referencia sólo a los sujetos que experimentaron mejoras partiendo de puntuaciones inicialmente bajas, salieron a la luz otros patrones de mediación, especialmente relevantes en la consecución de las mejoras obtenidas en aquellos procesos que más dificultad supusieron a los sujetos. De este modo, se produjo un incremento muy significativo en los patrones: *apoyo en la comprensión*; *recapitulación conjunta del trabajo realizado*; *repetición de una actividad con el objetivo de evaluar la puesta en práctica de un determinado proceso o de ampliar la información*; *mediación de la conciencia de lo que el sujeto sabe acerca del contenido de los textos trabajados*;

mediación de estrategias de supervisión; mediación de estrategias de autoevaluación; mediación de estrategias de transferencia y generalización; y mediación de la motivación, la actitud y el interés.

Con relación a la prueba de ajuste, los patrones de mediación más relevantes en el grupo experimental completo, fueron: el *apoyo en la comprensión*; la *mediación del uso de estrategias de supervisión*; y la *mediación de la motivación, la actitud y el interés*. En cambio, para el grupo de mejoradores, también se observó un incremento significativo precisamente con relación a los procesos que mayor dificultad representaron. De este modo, algunos de estos patrones pasaron a establecer correlaciones significativas con la puntuación dinámica de los procesos de integración texto-conocimientos previos: *profundización en la respuesta del alumno*; *conciencia sobre las propias estrategias*; *conciencia de lo que sabe...*; y *mediación de estrategias de autoevaluación*.

Con relación a la prueba de inteligencia y para el grupo completo, aquellos patrones especialmente relevantes fueron: la *profundización en la respuesta del alumno o en su justificación, haciendo nuevas preguntas o cuestionando sus respuestas* (que establece una correlación significativa con la mejora en CI); la *recapitulación conjunta del trabajo realizado*; la *repetición de una actividad con el objetivo de evaluar la puesta en práctica de un determinado proceso o de ampliar la información*; la *mediación de la conciencia sobre elementos estructurales*; la *mediación de la conciencia sobre las propias estrategias de comprensión*; así como *sobre lo que el sujeto sabe acerca del contenido de los textos trabajados*; la *mediación del uso de estrategias de planificación*; *mediación de estrategias de integración textual*; y *mediación de estrategias de integración del texto con conocimientos previos*. Cuando se tuvieron en cuenta sólo los sujetos mejoradores, se observaron incrementos generalizados en los niveles de correlación y, especialmente, con relación a los procesos que más dificultad representaron, pasando algunos de ellos a ser significativos estadísticamente. Los patrones más relevantes para este subgrupo fueron: *apoyo en la comprensión...*; *mediación de la conciencia de lo que el sujeto sabe acerca del contenido de los textos trabajados*; *mediación de estrategias de autoevaluación*; y *mediación de la motivación, la actitud y el interés*.

Por otra parte, uno de nuestros objetivos en el análisis de los resultados, era el establecimiento de las posibles diferencias en la mediación efectuada por los aplicadores en su puesta en práctica del dispositivo EDPL. En este sentido, pudimos comprobar, por una parte, la ausencia de relaciones significativas entre la cantidad de anotaciones relativas a los patrones de mediación en cada uno de los grupos de aplicación considerados, y la puntuación dinámica global obtenida. Por otra parte, las diferencias encontradas en el número de anotaciones, sí fueron claras y significativas con relación a los distintos grupos, así como con respecto a la casi totalidad de los patrones de mediación contemplados. De este modo, las notables diferencias observadas en la cantidad de anotaciones relativas a la mediación en cada uno de los grupos, no se reflejó de forma clara en las puntuaciones dinámicas globales obtenidas por cada uno de ellos. Si bien, en los contrastes efectuados para el conjunto del grupo experimental, sí parecía observarse una tendencia indicativa de una relación negativa entre la cantidad de mediación recibida y la puntuación dinámica obtenida, esta tendencia no tenía lugar de forma nítida en los análisis grupales, probablemente debido a que dentro de los grupos de aplicación, existían a su vez distintos niveles de competencia. La aplicación grupal del dispositivo dinámico en la mayoría de las ocasiones, ha conllevado la puesta en

práctica colectiva de patrones de mediación que, sin embargo, algunos sujetos podrían no haber necesitado para realizar las tareas. De esta forma, las anotaciones de la puesta en práctica de las conductas mediadoras no se habrían relacionado de forma adecuada con la competencia mostrada por cada uno de los sujetos en las actividades; competencia que, por otra parte, quedaba valorada igualmente por el aplicador mediante los indicadores de evaluación. Sin duda, un análisis más cualitativo sobre el contenido de los patrones metodológicos puestos en práctica que se han relacionado con las mejoras observadas, así como sobre el *modo* en que se ha concretado la mediación, y no tanto sobre la cantidad de la misma, puede ofrecernos información más valiosa acerca de la posible influencia de la mediación diferencial puesta en práctica por los distintos aplicadores.

Por último, con relación a los patrones metodológicos de evaluación dinámica utilizados en la mediación, hemos querido abordar de forma complementaria una serie de análisis encaminados a determinar la estructura factorial del conjunto de anotaciones relativas a la mediación. Estos análisis nos mostraron una estructura compuesta por once factores, que lograban explicar conjuntamente casi el 80% de la varianza. En este sentido, aunque la estructura del sistema de categorías presentado y utilizado, venía configurada por la propia estructura interna del dispositivo EDPL, conformándose en torno a los distintos procesos contemplados en la propuesta, hemos considerado oportuno contrastar nuestra categorización, realizada *a priori* y en función de nuestra teoría de base, con la que puede deducirse de análisis derivados exclusivamente de los datos proporcionados con base en las anotaciones y respuestas otorgadas. Los resultados de estos análisis nos han permitido agrupar, en función de las correlaciones establecidas, patrones de mediación dirigidos en un principio a facilitar la puesta en práctica de diferentes procesos implicados en la lectura. Esta nueva categorización nos ha ofrecido la posibilidad de observar cómo se han agrupado las diversas sugerencias y conductas de mediación llevadas a cabo por los aplicadores; o lo que es lo mismo, qué patrones metodológicos han sido puestos en práctica de forma conjunta. Aunque los criterios seguidos para la formulación de ambas categorizaciones, así como el sentido de los análisis que cabe efectuar, son distintos, la comparación entre ambas estructuras nos deja entrever importantes coincidencias, derivadas en su mayor parte de la consideración de los distintos procesos implicados en la lectura contemplados en el dispositivo.

Con relación a la estructura interna del dispositivo, hemos destacado el alto número de correlaciones significativas y muy significativas establecidas entre los distintos procesos contemplados, así como entre éstos y la puntuación dinámica global. Nos parece sumamente relevante la altísima correlación entre la puntuación dinámica global y la relativa a los procesos metacognitivos. Este dato apunta nítidamente a la relevancia que estos procesos han tenido en la puntuación EDPL y a la interrelación existente entre la valoración global sobre el proceso de comprensión y la valoración específica sobre los procesos metacognitivos. De este modo, la consistencia del dispositivo EDPL quedaría reforzada con relación a su contenido, a la teoría de base que lo sustenta, así como también al sistema de codificación y valoración utilizado. Los distintos procesos contemplados formarían de este modo parte de un mismo conjunto de elementos, que contribuirían a valorar el proceso de comprensión. En este sentido, podemos concluir que los procesos implicados en la lectura que han sido contemplados en nuestra investigación, contribuyen significativamente a la evaluación global de la lectura y la comprensión, aún cuando quede patente su entidad y contribución diferencial al análisis.

Por otra parte, hemos expuesto y analizado el conjunto de actividades contempladas, su estructura interna y su disposición dentro de la galaxia propuesta. Así mismo, hemos profundizado en la propuesta de actividades contenida en cada uno de los bloques de procesos contemplados en el dispositivo. En cuanto a su materialización en la aplicación experimental, como apuntábamos en la narración del procedimiento descrito, la puesta en práctica de todas las actividades del dispositivo que habían sido previstas para este estudio, no pudo finalmente llevarse a cabo por diversos problemas, relativos sobre todo al tiempo del que los aplicadores disponían realmente para esta labor o a las dificultades añadidas que supuso la aplicación del dispositivo a pequeños grupos y no de forma individual. Como consecuencia lógica, la propuesta de intervención centrada en los procesos metacognitivos, no pudo ser trabajada por ninguno de los aplicadores. Estas circunstancias no habían sido imprevistas en absoluto por nosotros, ya que, en función de las características de nuestra propuesta, contábamos con la posibilidad de que la aplicación del dispositivo, en la forma en que era propuesta para su aplicación experimental en este estudio, no pudiera completarse debido al número de actividades que se contemplaban y al tiempo que podría ser necesario para poner en práctica la posterior intervención; esto al margen de la consideración de los avatares propios de la aplicación experimental de un proceso educativo en el contexto escolar.

Conexión de la aplicación experimental con nuestra teoría de base y mecanismos de acción del dispositivo EDPL

El dispositivo de evaluación dinámica EDPL, que ha sido aplicado experimentalmente, está orientado especialmente a la valoración de los procesos metacognitivos implicados en la comprensión de textos. Como quedaba recogido en nuestro marco teórico, son precisamente estos procesos los que admitirían en mayor medida una acción educativa que se mostrara eficaz para optimizar el funcionamiento cognitivo del alumnado con dificultades de aprendizaje o discapacidad intelectual (Vygotski, 1931/1997; 1934/1995; Martín y Marchessi, 1990; Luria, 1995; Mora, 1998; Carlson y Wiedl, 2000; Navarro y Mora, 2002). Hemos tenido la oportunidad de revisar teóricamente, y de comprobar así mismo, que entre las carencias más significativas que presentaban los sujetos con dificultades de aprendizaje o déficit cognitivo, aquellas relacionadas con los procesos de control y regulación consciente de los propios procesos de acción o pensamiento, ocupaban un lugar destacado. De igual forma, también recogíamos que, precisamente, los sujetos con problemas en su aprendizaje parecían beneficiarse de manera especial de una intervención centrada preferentemente en los procesos de autorregulación (Meichenbaum, 1986; Kurtz y Borkowski, 1987; Campione, Brown y Ferrara, 1987; Scrugg y Mastropieri, 1992; Mastropieri y Scrugg, 1997; Mora, 1998; Saldaña, 2001).

Por otra parte, nuestro dispositivo EDPL fue pensado y desarrollado con el objetivo de optimizar la evaluación de los procesos implicados en la lectura, especialmente con relación a aquellos sujetos que presenten dificultades relacionadas con la lectura; recabando información que pudiera resultar valiosa para orientar hacia la mejora el proceso de enseñanza-aprendizaje. Esta información valiosa tiene mucho que ver, por una parte, con los procesos de interacción mediada que tienen lugar en el transcurso de la aplicación de las actividades propuestas en el dispositivo, extrayendo los aplicadores información acerca de aquellos patrones metodológicos de evaluación dinámica utilizados en la mediación, que provocaron una puesta en práctica de habilidades de comprensión de mayor calidad; así como también tiene que ver, por otra parte, con la

valoración preferente de las habilidades metacognitivas del sujeto en el desarrollo de las tareas, como elemento clave para la obtención de información que permita implementar acciones educativas que conduzcan a mejoras significativas y duraderas.

Por otra parte, en la valoración efectuada por los aplicadores acerca del dispositivo de evaluación dinámica, y de su aplicación experimental, a través del cuestionario ya citado (anexo X), se recogían algunos aspectos considerados *clave* para el éxito de EDPL y la obtención de información valiosa con relación a una posterior intervención. En este sentido, los aplicadores hicieron especial mención de elementos relacionados con la propuesta metodológica realizada, es decir, de elementos como el propio procedimiento de evaluación dinámica, la mediación de procesos, centrarse en los indicadores de evaluación o la puesta en práctica de los patrones de mediación. Así mismo, se destacaron otros aspectos relacionados con el registro de la información, el tipo de procesos contemplados en el dispositivo, el trabajo colectivo realizado por los propios aplicadores o la propuesta de actividades contenida en el Manual.

Con relación a las razones de las mejoras observadas, es decir, a aquellos elementos del proceso considerados *clave* para explicar el cambio en los sujetos experimentales, los propios aplicadores consideraron al mismo nivel de importancia los elementos relacionados con la propuesta metodológica realizada (mediación de procesos metacognitivos, lentificación del proceso en aquellos bloques de actividades donde se hallaron dificultades, puesta en práctica del dispositivo de forma lúdica y participativa, etc.), y los elementos relacionados con los procesos trabajados en el dispositivo (estrategias de aprendizaje trabajadas, trabajo sistemático sobre habilidades metacognitivas, intervención sobre distintos procesos implicados en la lectura, etc.). Así mismo, también pusieron de manifiesto la relevancia de los elementos relacionados con el esquema de actividades propuesto en EDPL (secuenciación y programación de las actividades, seguimiento de las pautas metodológicas propuestas, etc.). Con relación al trabajo sistemático sobre las habilidades metacognitivas, Carlson y Wiedl (2000), así como Cormier, Carlson y Das (1990), o Kar et al. (1993), encontraron que en la reestructuración de las pruebas cognitivas, en el sentido de la introducción de retroalimentación detallada y de verbalización antes, durante y después de la tarea, como factores clave para la mejora de los sujetos; eran precisamente aquellos con dificultades relacionadas con los procesos metacognitivos, como la planificación, los que mejoraban significativamente su rendimiento. La mayor eficacia del modelo de evaluación propuesto por estos autores, residía de esta forma en la valoración de las habilidades metacognitivas en los sujetos que presentaban déficit en este tipo de procesos. Apuntábamos cuando analizábamos este modelo, que la relevancia que tendrían para la mejora, los procesos de verbalización, antes, durante y después de la actividad, podría estar relacionada con una posible reestructuración de los procesos de pensamiento, en el sentido de optimizar los recursos atencionales, lo que podría afectar positivamente a la consideración de las reglas que deben ser seguidas, las estrategias que deben ser empleadas o la información que debe ser recordada. Por otra parte, un segundo factor del proceso de verbalización, que puede incidir notablemente en la mejora de la ejecución de las tareas, es que éste supone en la práctica el uso de un sistema de codificación adicional; en efecto, cuando se introduce la verbalización del proceso de aprendizaje, puede llegar a producirse una codificación doble y conjunta de la información, lo que redundará en un proceso de aprendizaje con mayores garantías. En nuestro caso, las mejoras postratamiento observadas en los resultados de la prueba de inteligencia general, podrían ser explicadas desde esta hipótesis. De este modo, los

sujetos que obtuvieron ganancias posttest podrían haber utilizado inicialmente un sistema de codificación preferentemente perceptivo-visual, el cual se habría revelado claramente insuficiente para dar cuenta con cierto éxito de la tarea a realizar. Posteriormente a la implementación del tratamiento, habría tenido lugar una conjunción de dos sistemas de codificación de la información; por una parte, el sistema perceptivo-visual, que ya estaba presente; y por otro, un sistema de codificación simbólico-lingüístico, fruto de la apropiación de estrategias de autorregulación verbal mediadas en el transcurso de la aplicación del dispositivo dinámico.

Nuestro dispositivo de ED se habría dirigido fundamentalmente a las dos dimensiones consideradas por Carlson y Wiedl, dentro de su modelo de inteligencia (Carlson y Wiedl, 2000), como aquellas más fructíferas en la aplicación de procesos de evaluación dinámica, así como en la validación de los mismos. Estos autores destacan, por una parte, las *estructuras cognitivas generales*, que harían referencia básicamente a las habilidades metacognitivas de autorregulación, y que se encontrarían estrechamente relacionadas con los logros obtenidos en la otra de las dimensiones contempladas, es decir, en las *estructuras cognitivas específicas*, con las que hacen referencia a los dominios y funciones específicas, como la lectura, la aritmética o la memoria de trabajo. El dispositivo EDPL actuaría sobre estas dos dimensiones, aunque también pretende incidir de forma significativa sobre el contexto escolar y los procesos de ajuste personal-social, entre los que considera el proceso de enseñanza y aprendizaje. Por otra parte, los resultados obtenidos no descartan que pudiera ejercer influencia sobre otra de las dimensiones apuntada por estos autores; esta es la dimensión *inteligencia fluida-factor «G»*. Esta influencia la habría ejercido, como hemos tenido oportunidad de comprobar, a través de los procesos metacognitivos. En este sentido, las investigaciones llevadas a cabo por Carlson y Wiedl, señalaban tres pasos necesarios con vistas a lograr una optimización de los esfuerzos dirigidos a la implementación de procesos de ED. El primero de ellos hacía referencia al establecimiento teórico y empírico de las relaciones entre las estructuras cognitivas generales y los dominios/funciones específicos; el segundo destacaba la necesidad de demostrar la validez predictiva de las pruebas dinámicas sobre los dominios específicos; y el tercero hablaba de desarrollar procesos de intervención educativa basados en la teoría de las estructuras cognitivas generales (acción preferente sobre los procesos metacognitivos), determinando además la mayor validez predictiva de las pruebas dinámicas sobre una evaluación de corte tradicional. En nuestro caso, el análisis de las bases teóricas y metodológicas, así como del proceso y los resultados de nuestra investigación, nos ofrece la posibilidad de establecer vínculos teóricos y empíricos entre las estructuras cognitivas generales y los dominios específicos, por cuanto los resultados muestran la relación entre la puesta en práctica de habilidades de autorregulación y el rendimiento en otras funciones específicas (mediante las valoraciones del profesorado), así como en la comprensión de textos. Del mismo modo, hemos mostrado a raíz del análisis de los resultados, que las puntuaciones en EDPL predicen de forma significativa el rendimiento en lectura. Así mismo, nuestro dispositivo está desarrollado sobre la base de la relevancia de mediar el uso de los procesos metacognitivos en el desarrollo de estrategias de aprendizaje. Por último, hemos determinado el mayor valor predictivo de las puntuaciones dinámicas sobre aproximaciones estáticas de evaluación de la lectura.

El modelo de Prueba de los límites, propuesto por Carlson y Wiedl (2000), ha mostrado la idoneidad de intervenir sobre procesos de autorregulación y conocimiento metacognitivo para mejorar el rendimiento de aquellos sujetos que muestran

precisamente un bajo nivel de funcionamiento metacognitivo. Por otra parte, de sus estudios también se desprende la importancia de que la metodología de evaluación adoptada, así como la metodología de enseñanza, guarden estrecha relación. En este sentido, tendrían especial relevancia los métodos de enseñanza acordes con los principios de la enseñanza adaptativa y, en general, con una concepción constructivista del proceso de enseñanza y aprendizaje. Nuestros propios resultados confirman y refuerzan los obtenidos por Carlson y Wiedl; y, por otra parte, nuestra propuesta metodológica para la aplicación contextualizada del dispositivo EDPL, se fundamenta en los principios mencionados más arriba. En definitiva, los resultados obtenidos por estos autores, y por nosotros mismos, con relación a la opción preferente por la mediación de los procesos metacognitivos; así como las propuestas metodológicas que se desprenden de los estudios realizados, apuntan de esta forma en la misma dirección que los de la mayoría de los autores de procedimientos dinámicos de evaluación, comenzando por el propio Vygotski (Vygotski, 1934/1995; 1931/1997; Campione y Brown, 1987; Campione, Brown y Ferrara, 1987; Jensen, 2000; Resing, 2000; Saldaña, 2001; Sternberg y Grigorenko, 2003; Navarro y Mora, 2003b; Navarro, 2004).

10.3 Discusión del proceso de evaluación seguido para valorar la aplicación experimental del dispositivo EDPL: metaevaluación

En capítulos anteriores de nuestro estudio hemos realizado un análisis de cada uno de los aspectos considerados relevantes en la valoración de un procedimiento dinámico de evaluación. Hemos aportado datos cuantitativos y cualitativos que permiten valorar elementos relativos a la validez y la fiabilidad, tanto de las pruebas que han sido utilizadas como criterio para la medición cuantitativa del cambio, como del dispositivo EDPL; así mismo, estos datos permiten evaluar la calidad de la información aportada por el dispositivo dinámico, a través de su implementación experimental en los centros. Por otra parte, una cierta peculiaridad de nuestro estudio la constituye el hecho de que lo que se está aplicando experimentalmente es un proceso de evaluación en sí mismo, por lo que la valoración de esta aplicación se realiza a su vez sobre un procedimiento de análisis y evaluación. Con relación a este análisis sobre nuestro propio proceso de evaluación de la aplicación experimental del dispositivo, a lo largo del desarrollo de la investigación, hemos tratado de justificar en lo posible cada una de las decisiones tomadas que tenían que ver con los procedimientos seguidos para analizar los posibles cambios en el grupo experimental. En este sentido, en el capítulo dedicado al diseño de la investigación, hemos expuesto y razonado la opción adoptada relativa a la estructura del diseño, es decir, pretest-postest con grupo de control no equivalente. Así mismo, hemos analizado las limitaciones del diseño, así como las técnicas de control adoptadas para minimizar las posibles fuentes de distorsión de los resultados. También hemos hecho referencia a la selección diferencial de los sujetos participantes del estudio, con base en su pertenencia al grupo de sujetos que presentaban especiales dificultades relacionadas con los procesos de comprensión. Esta opción se plantea en nuestra investigación como la mejor posible, en función de los objetivos del estudio y de algunas de las hipótesis formuladas. Así mismo, la consideración de distintos subgrupos en la muestra y en nuestro diseño de investigación, relativa a la presencia de circunstancias o características de índole cognitiva o socio-cultural, nos ha permitido acceder a la valoración diferencial de los efectos de la aplicación del dispositivo, en función de la pertenencia de los sujetos a estos subgrupos contemplados.

Por otra parte, la utilización de diversas pruebas criterio ha permitido acceder a una información más compleja y valiosa sobre los procesos analizados. Unida a esta información, proporcionada por los resultados de las pruebas, se encuentra la que hemos podido obtener de la valoración del proceso y de los resultados de la aplicación del dispositivo dinámico, así como a través de la cumplimentación, por parte del profesorado colaborador, de los cuestionarios de valoración del dispositivo EDPL. De esta forma, se ha producido una cierta triangulación de la información obtenida, que ha contribuido a incrementar el proceso de control experimental, por cuanto ha sido posible analizar comparativamente y de modo complementario, los datos provenientes de diversas fuentes de carácter cuantitativo y cualitativo. En este sentido, se argumentaba en el análisis del diseño de investigación, que la confluencia de los datos en una dirección acorde a la apuntada por nuestras hipótesis, aportaría un mayor grado de validez a los resultados obtenidos. Así mismo, se hacía referencia a la posibilidad de enriquecimiento del análisis, también ante la obtención de resultados divergentes, en el sentido de ofrecer información complementaria sobre distintos grados de calidad de respuesta al tratamiento, otorgados en función de condiciones o estilos docentes distintos.

La bondad del diseño utilizado se ha visto reforzada por la adopción de algunas medidas, tales como tomar a los sujetos de los grupos experimental y control de la misma población, es decir, del grupo de sujetos que presenta especiales dificultades de aprendizaje relacionadas con la lectura, y que, así mismo, son atendidos mayoritariamente en las aulas de apoyo a la integración y/o de audición y lenguaje. Estos sujetos pertenecían además a centros educativos con características socio-culturales similares (nivel medio-bajo), y, en ocasiones, a los mismos centros. Del mismo modo, hemos hecho alusión a que los profesores colaboradores que participaron, formando parte del grupo experimental o del grupo control, pertenecían a grupos de docentes comprometidos en proyectos educativos de innovación en sus respectivos centros. La asignación final de los sujetos al grupo experimental o control ha estado en función de factores relativos a las posibilidades de participación en las sesiones del Grupo de Trabajo en el que se formó a los aplicadores del dispositivo y tuvo lugar el análisis de su puesta en práctica. Todos los subgrupos contemplados (alumnado con dificultades de aprendizaje de la lectura en ausencia de discapacidad o privación; sujetos con discapacidad cognitiva; sujetos con privación socio-cultural; y alumnado que recibía apoyo a la integración), estaban representados en cada una de las condiciones experimentales, lo que ha contribuido a incrementar la validez externa del estudio. El análisis de los resultados se ha realizado en función de los diversos subgrupos que conformaban la muestra experimental. En este sentido, la posibilidad de realizar este análisis ha comportado una doble ventaja. Por un lado, nos ha permitido delimitar mejor el ámbito de eficacia del dispositivo EDPL, así como las condiciones que optimizarían su aplicación, al menos en función de la aplicación experimental realizada; por otra parte, permite orientar diferencialmente la intervención en función del análisis del proceso y de los resultados para cada uno de los grupos.

La estructura del diseño utilizado nos ha permitido controlar las principales fuentes de error en un estudio como el nuestro, tales como la edad de los sujetos, el nivel escolar, la inteligencia, el nivel de ajuste personal-social o el nivel de rendimiento previo en lectura. Los niveles de equivalencia obtenidos en los diversos contrastes realizados son más que aceptables, habiendo procedido a efectuar análisis diferenciales para algunos subgrupos cuando así lo aconsejaban los resultados de los citados contrastes. Las

puntuaciones de los grupos experimental y control completos eran equivalentes en el pretest, tanto en nivel escolar, ajuste personal-social, comprensión y nivel de inteligencia. Estos análisis han contribuido a ejercer un control sobre los efectos principales de regresión a la media, que, de otro modo, hubieran puesto en cuestión las ganancias observadas por el grupo de sujetos que partía de puntuaciones iniciales más bajas.

Cuando se procedió a realizar contrastes en función del alumnado que recibía apoyo a la integración, encontramos que el número de sujetos que participaba de este tipo de apoyo en el grupo experimental, era significativamente más alto que el que lo hacía en el grupo control. Estas diferencias tenían lugar sobre todo en primaria, lo que podría explicar en parte las mayores puntuaciones posttest obtenidas por el grupo de control. En este mismo sentido, en los casos en los que los sujetos no recibían apoyo a la integración, ha sido más difícil determinar si se cumplían las características esenciales requeridas en el estudio. De este modo, parte del alumnado seleccionado en primera instancia, lo habría sido sólo con base en la información proporcionada por los tutores a los profesores colaboradores; por lo que un número importante de integrantes del grupo control, aún habiendo sido catalogados como sujetos con dificultades de aprendizaje en la lectura, guardarían notables diferencias de partida con los integrantes del grupo experimental. En definitiva, las comparaciones efectuadas en función del primero de los criterios adoptados, podrían contener un cierto sesgo derivado de la selección inicial de los sujetos experimentales adscritos a uno de los grupos. En este sentido, dado que hablamos de sujetos que no recibían apoyo especializado en las aulas de apoyo a la integración o de audición y lenguaje, ámbitos en los que desarrollaban su trabajo los profesores-aplicadores del dispositivo, éstos pasaron a engrosar las filas del grupo control.

Las puntuaciones son asimismo equivalentes en cuanto a su dispersión y no sólo con relación a su promedio. De esta forma, el nivel de equivalencia real entre los grupos experimental y control ha sido mayor. Igualmente, hemos comprobado que existe homogeneidad de varianzas para los grupos experimental y control completos, así como para la mayoría de los contrastes efectuados entre los distintos subgrupos contemplados. Por último, además de la estructura del diseño, la composición de los grupos de tratamiento y control, que, como hemos visto estaban formados por sujetos de distintas edades y niveles educativos, permitía controlar en cierto modo los posibles efectos de la maduración.

Otros aspectos recogidos en procesos de análisis sobre los procedimientos de evaluación, seguidos en estudios que comparten algunas características relevantes con el nuestro, están relacionados, por una parte, con la consideración de grupos de aplicación con distinto nivel de calidad y profundidad en su implementación del tratamiento experimental; así como, por otra, con la utilización de los profesores habituales de los sujetos experimentales para aplicar el dispositivo. En ambos casos se ha descrito que se produce un incremento de la validez externa (Campbell y Stanley, 1978, citado en Mora, 1998). En el primero de los casos, la ventaja se produce por el hecho de que las hipótesis experimentales se contrastan —y verifican en su caso— en diversas condiciones de calidad de aplicación. De obtenerse resultados positivos en todos los casos, se mostraría una mayor potencia del tratamiento, que sería capaz de actuar con distintos profesores y ante circunstancias y condiciones también distintas. En el segundo caso, las razones tienen que ver con la mayor validez que presentaría el tratamiento para

el contexto natural para el que ha sido pensado y desarrollado. Esta mayor validez se relaciona con el hecho de que sean los propios profesores los que aplican el dispositivo, ya que esto mostraría que es posible para el profesorado aprender el procedimiento y aplicarlo con éxito. Nuestros resultados muestran que se han llegado a producir mejoras significativas generalizadas a varios de los grupos de aplicación; grupos que, por otra parte, desarrollaron la puesta en práctica del dispositivo bajo circunstancias diversas. Estas mejoras han tenido lugar, como hemos podido comprobar, de forma desigual en función de la tarea criterio que fue utilizada; del criterio de selección tenido en cuenta para realizar los análisis; o del subgrupo de pertenencia de los sujetos con relación a sus características cognitivas o socio-culturales. Por otra parte, el análisis del proceso y de los resultados obtenidos, avalan notablemente el aprendizaje efectuado por los profesores-aplicadores sobre el procedimiento dinámico, así como su aplicación efectiva en el contexto escolar en el que desarrollaban su labor docente.

Sin ninguna duda, podemos afirmar que todas aquellas situaciones que, de un modo u otro, han dificultado la aplicación del dispositivo tal y como fue diseñada en nuestra investigación, alejándose por tanto de las condiciones óptimas para la implementación de un procedimiento de evaluación dinámica como el nuestro, han jugado en contra de nuestras hipótesis experimentales. En este sentido, debemos destacar las dificultades de tipo institucional-organizativo que presidieron la aplicación de EDPL en algunos de los centros participantes. En uno de los centros, las actividades propuestas para el GE dieron comienzo más tarde que en el resto, a lo que debemos unir la existencia de algunos problemas para realizar una aplicación sistemática y continuada del dispositivo; problemas derivados de la necesidad, en función de la responsabilidad de la aplicadora como parte del Equipo directivo del centro, de atender a menudo asuntos de carácter organizativo. Por otra parte, en este mismo centro, así como en otros de la muestra, tuvieron lugar determinadas circunstancias relativas al elevado absentismo de algunos sujetos que formaban parte de la muestra experimental. Otra de las circunstancias a las que hicimos referencia cuando analizábamos el proceso de aplicación de EDPL, tuvo que ver con la necesidad de aplicación colectiva del dispositivo, contando con un elevado número de sujetos en algunos de los grupos. En todos estos casos, las circunstancias contextuales de aplicación han actuado de modo conservador, reduciendo el tiempo y las condiciones óptimas de aplicación en las cuales debía evaluarse EDPL. Algunas de estas circunstancias podrían haber tenido un peso significativo en la ausencia de mejores resultados en algunos de los grupos de aplicación. En este sentido, cabe recordar la cita que hacíamos de Meloth (1990) en nuestro marco teórico, en la que se apuntaba la posibilidad de que un periodo de intervención «corto» no fuera del todo suficiente para reflejar ganancias significativas en el uso de estrategias metacognitivas, siendo tal vez necesario un espacio de tiempo mínimo (16 semanas) para observar efectos significativos, sobre todo con niños y niñas que presentan dificultades de aprendizaje. En nuestro caso, el tiempo efectivo de aplicación del dispositivo para la mayoría de los grupos de aplicación, ha rondado esas 16 semanas; sin embargo, somos conscientes de que en algunos casos, los problemas que hemos mencionado anteriormente, han reducido sensiblemente este periodo. Sin embargo, a pesar de los datos expuestos anteriormente, se han obtenido ganancias significativas en la dirección apuntada por nuestras hipótesis, al tiempo que se ha constatado una valoración especialmente positiva del profesorado colaborador. En este sentido, se refuerzan notablemente las hipótesis planteadas y se fortalecen nuestras convicciones relativas a la idoneidad de implementar en el contexto escolar procedimientos dinámicos de evaluación de los procesos implicados en la lectura.

En cuanto a la batería de pruebas utilizada como criterio, ya expusimos los motivos que nos llevaron a elaborar dos de las tres pruebas aplicadas. Éstas han tratado de responder lo más ampliamente posible a la necesidad de evaluar, por una parte, el rendimiento en lectura, tomando en consideración los distintos procesos implicados en la comprensión de los textos; y, por otra, el nivel de ajuste personal-social. Con relación a la prueba para evaluar el nivel de inteligencia, hemos utilizado el test de factor «G» de Cattell-2 en su versión A, de sobrado reconocimiento para ofrecer una medida del rendimiento cognitivo con niveles altos de fiabilidad y validez.

La elaboración y aplicación de pruebas de criterio construidas por nosotros mismos, llevaba consigo la necesidad de evaluar exhaustivamente distintos aspectos relacionados con la validez y la fiabilidad. En este sentido, con relación a la validez de la escala APSL, debemos hacer constar que los índices de dificultad de la prueba se incrementaron significativamente cuando se analizaron los resultados del alumnado que presentaba dificultades de aprendizaje; por otra parte, el dato anterior nos hablaría de la validez de constructo de la escala, ya que era esperable que los sujetos con dificultades de aprendizaje en la lectura obtuvieran puntuaciones más bajas en la escala. Los contrastes realizados entre los resultados obtenidos por la muestra de aplicación piloto y los sujetos con dificultades que conformaron los grupos experimental y control, indicaron diferencias muy significativas a favor del grupo piloto, constituyendo este último una referencia crucial para los intereses de nuestro estudio. Los índices de fiabilidad y de homogeneidad ofrecieron garantías sobre la validez de la escala con relación a la evaluación del ajuste personal-social. La consistencia interna de las puntuaciones y la inter-relación de las distintas dimensiones, nos hablan sobre la unidad que presenta la prueba en cuanto a los bloques evaluados. La puntuación en una dimensión informa y explica en parte la puntuación en las demás dimensiones y en la puntuación global. En este sentido, todas las dimensiones mantienen niveles de correlación muy significativos con la puntuación global. Con relación a la validez externa, hemos de apuntar que la designación previa del profesorado y de los orientadores acerca de la presencia de dificultades de aprendizaje de la lectura, quedaba avalada en función de las diferencias muy significativas entre las puntuaciones obtenidas por los sujetos que presentaban dificultades y aquellos que no presentaban problemas de aprendizaje. La correlación entre la puntuación global y las dimensiones metacognitivas es muy significativa, poniendo de manifiesto la relevancia de estas dimensiones en la explicación de los resultados obtenidos, al tiempo que refuerza la relación existente entre el ajuste personal-social y la metacognición (Mora, 1998; Mora-Merchán y Mora, 2000). Por último, reseñar también la validación efectuada por el grupo de expertos, realizada previamente a la aplicación de la escala y en la que ya profundizamos cuando analizamos pormenorizadamente el instrumento. En general, podemos decir que la valoración efectuada sobre la adecuación de la escala para evaluar los niveles de ajuste personal-social con relación a la lectura, fueron muy positivos.

Con relación a la prueba ECO 1, también debemos destacar que los índices de dificultad se incrementaron muy significativamente cuando fueron analizados los resultados del grupo de sujetos con dificultades de aprendizaje en su lectura, tanto en el grupo experimental como en el grupo control, con respecto a los obtenidos por la muestra piloto. Este resultado incrementa la validez de constructo de la prueba elaborada. Los indicadores de fiabilidad para los ítems y para las dimensiones, y los índices de homogeneidad, ofrecen garantías sobre la validez interna de la prueba. Los índices de homogeneidad muestran que cada uno de los ítems de la prueba contribuye de manera

significativa a evaluar lo mismo que el conjunto de elementos que la conforman. La matriz de correlaciones muestra una fuerte conexión entre las puntuaciones obtenidas en las distintas dimensiones contempladas en la prueba; en este sentido, todas las correlaciones son muy significativas, estableciendo, así mismo, la puntuación total una correlación muy significativa con cada una de las dimensiones de la prueba. De esta forma, la calidad de la comprensión de textos mostrada en cada una de las dimensiones, se reflejaría de forma significativa en la calidad de la comprensión en el resto de dimensiones, así como en la puntuación total. La puntuación en una dimensión explica en parte la obtenida en las otras. Por otra parte, la validación realizada por el grupo de expertos otorgó en general una valoración muy positiva a la prueba ECO 1. Esta valoración se realizaba sobre la adecuación de la prueba para evaluar la comprensión de textos en educación primaria. Por último, con relación también a la validez externa, podemos afirmar que la designación previa del alumnado con especiales dificultades de aprendizaje de la lectura, por parte del profesorado y de los orientadores, se vio refrendada por los resultados diferenciales de los grupos de sujetos con especiales dificultades y sin ellas. Si bien apuntábamos anteriormente esta designación previa del profesorado sobre la presencia de dificultades de aprendizaje, como una de las posibles causas de las diferencias observadas entre los grupos experimental y control, diferencias en cuanto al número de sujetos que recibían apoyos más específicos; los resultados de los contrastes efectuados relacionan significativamente la designación y las puntuaciones obtenidas por el alumnado con y sin dificultades de aprendizaje. Aunque podría entenderse esto como una interpretación contradictoria, quizás una sencilla explicación pueda encontrarse en el peso diferencial que pueden haber tenido en las correlaciones con la designación efectuada por el profesorado, aquellos alumnos que recibían apoyo a la integración, dentro del grupo de sujetos con dificultades. Por último, los niveles de correlación entre la puntuación total y las dimensiones metacognitivas, revelan la importancia de estas dimensiones en la explicación del resultado obtenido por los sujetos en la prueba. De esta forma, se ve reforzada la relación entre el contenido evaluado en estas dimensiones y la comprensión.

Otro de los aspectos relevantes con relación a la prueba ECO 1, que fue ya destacado en la validación de expertos y analizado en el capítulo de Instrumentos, lo constituye la excesiva dificultad que podría haber representado la prueba para el alumnado de primaria con especiales dificultades de lectura. El índice de dificultad de la prueba se situaba bastante por encima del nivel medio ya en la muestra piloto, compuesta de forma mayoritaria por alumnado sin especiales dificultades de aprendizaje. En efecto, comentábamos que podría haber tenido lugar un cierto efecto *techo* por el cual algunos sujetos, a pesar de haber obtenido mejoras importantes —a tenor de las valoraciones cualitativas efectuadas por los aplicadores— en aspectos relacionados con las habilidades trabajadas en el dispositivo, no habrían conseguido ganancias en la realización postratamiento de la prueba. Esto habría ocurrido a pesar de los buenos indicadores estadísticos en fiabilidad, homogeneidad y correlación entre las distintas dimensiones contempladas. Debemos recordar que bajo el segundo y el tercero de los criterios de análisis de resultados, los sujetos del grupo experimental obtuvieron ganancias significativas en el postest. Esto hace indicar que fueron precisamente los sujetos que presentaban más dificultades, los que podrían haber obtenido un mayor beneficio de su participación en la aplicación del dispositivo, obteniendo mejoras muy significativas en esta prueba criterio. En cambio, fueron los sujetos que partían de puntuaciones un poco más altas, es decir, por encima de los ochenta y cuatro puntos, los que no mostraron mejoras postest, y probablemente, aquellos para los que pudo tener

lugar el efecto techo debido a la dificultad de la prueba. En este sentido, el tiempo de aplicación no habría sido suficiente para lograr mejoras significativas en esta prueba para este subgrupo. Ya apuntamos que la composición de los grupos en experimental y control pudo no ser equivalente del todo. El grupo control habría contado con un cuarto subgrupo, además de sujetos con dificultades de aprendizaje, discapacidad y privación; este cuarto subgrupo habría estado compuesto por sujetos con algunos problemas de aprendizaje relacionados con la lectura, pero que no recibían ningún tipo de apoyo y que, posiblemente, tendrían un nivel de competencia mayor que los sujetos que sí recibían apoyo específico. La comprobación que realizamos de esta posibilidad nos mostró diferencias casi significativas para la comparación relativa a los sujetos que recibían apoyo a la integración, y diferencias muy significativas para el contraste relativo a los sujetos con privación.

Con relación a la prueba de comprensión elaborada para secundaria (ECOS), los índices de dificultad, al contrario de lo que ocurría en la prueba de primaria, muestran que la realización de la prueba no ha representado un nivel de dificultad que pueda considerarse excesivo. Por otra parte, al igual que ya ocurriera con la escala de ajuste y la prueba de primaria, los índices de dificultad también se vieron incrementados significativamente para el grupo de sujetos con especiales dificultades en la lectura. Este dato confirmaba nuestra hipótesis y reforzaba la validez de la prueba ECOS, por cuanto el alumnado con dificultades de aprendizaje obtenía significativamente menores puntuaciones en una medida de rendimiento en lectura. Así mismo, esto contribuye a validar externamente la prueba, ya que los resultados convergen con la valoración del profesorado en la designación del alumnado que presentaba especiales dificultades en lectura. Por su parte, los índices de homogeneidad promedio, así como los establecidos para cada ítem, muestran que la gran mayoría de los ítems de la prueba contribuyen de manera significativa a evaluar lo mismo que la prueba en su conjunto. La mayoría de las correlaciones entre las distintas dimensiones contempladas en la prueba presentan significación estadística, y la puntuación total obtenida correlaciona muy significativamente con todas las dimensiones contempladas. Los datos revelan una fuerte conexión entre los distintos procesos evaluados en la prueba, lo que a su vez informa sobre la unidad que presenta ésta con relación a los diversos bloques de contenido considerados. De esta forma, la puntuación obtenida en cada una de las dimensiones explicaría en parte la puntuación en las otras. Por otra parte, el índice de fiabilidad para el conjunto de ítems de la prueba es muy alto y, así mismo, el índice de fiabilidad para las dimensiones.

Por último, al igual que ocurre en la escala de ajuste y en la prueba de comprensión de primaria, los niveles de correlación hallados entre la puntuación global y las dimensiones metacognitivas, es decir, la dimensión de metaconocimiento estratégico y la de supervisión y autoevaluación del proceso de comprensión, así como la de concepción sobre la lectura, apuntan a su relevancia final en el resultado obtenido por los sujetos en la prueba, reforzando así mismo la relación entre el contenido evaluado en estas dimensiones y la comprensión efectiva de los textos.

Por lo que respecta a la evaluación del nivel de inteligencia general, ésta iba a tener lugar en un principio sólo en la fase inicial del estudio, mediante la aplicación de una prueba de CI. Contemplábamos esta opción en virtud de la relevancia que podría tener establecer la equivalencia inicial entre los grupos del estudio en una prueba de factor «G». Por otra parte, ya expusimos en el capítulo dedicado al diseño de la investigación,

que el tema central del estudio lo constituía la valoración de los procesos de comprensión, por lo que valorar las puntuaciones de Cattell en el postest no parecía del todo necesario. Fue posteriormente, a lo largo del proceso de investigación, aunque con anterioridad a la puesta en marcha de la fase experimental, cuando se vio oportuno evaluar el rendimiento cognitivo del alumnado una vez aplicado el dispositivo EDPL. La hipótesis que se revelaba cada vez con mayor claridad estaba fundamentada en que el trabajo desarrollado con la aplicación del dispositivo, se dirigía especialmente a los procesos metacognitivos, es decir, conocimiento metacognitivo y procesos de control y regulación de los propios aprendizajes, lo que suponía en la práctica un trabajo sobre habilidades y estrategias de ámbito general, susceptibles de ser aplicadas a otras actividades, aunque no tuvieran un contenido específico relacionado con los procesos de lectura. En efecto, de este modo, la intervención específica sobre los procesos metacognitivos, característica esencial de la aplicación del dispositivo de evaluación dinámica EDPL, habría tenido una repercusión sobre el rendimiento en la prueba de factor «G». Esta repercusión, en respuesta a lo expresado en nuestra tercera hipótesis, se habría producido como consecuencia de un efecto de transferencia, ya que el trabajo sobre metacognición se ha desarrollado sobre procesos implicados en un ámbito específico como la comprensión de textos. Los sujetos habrían aplicado en el postest habilidades y estrategias generales incorporadas e interiorizadas durante su participación en las sesiones de trabajo del dispositivo. En este sentido, tal y como pudimos comprobar al revisar los estudios de Resing (Resing, 1993, 2000), el componente metacognitivo de la metodología aplicada en la fase de entrenamiento del LIR, habría sido el elemento determinante en la generalización de los resultados que indicaban mayores índices de capacidad de aprendizaje. Por tanto, apuntábamos nosotros, cabía esperar que fueran estos componentes más estratégicos y generales, en mayor medida que otras habilidades más específicas, como el razonamiento inductivo, los que habrían contribuido de forma significativa a que la interiorización de estrategias de control y regulación cognitiva empleadas por el evaluador en la fase de las sugerencias graduadas (en nuestro caso, durante la mediación de los procesos metacognitivos), pasaran a formar parte progresivamente del repertorio estratégico y metacognitivo del sujeto, permitiéndole actuar de forma igualmente estratégica en la resolución de otras tareas, independientemente de su contenido. En nuestro caso, tal y como hemos apuntado, habría sido el componente metacognitivo del dispositivo EDPL, junto al uso de tareas contextualizadas, el que habría facilitado y permitido la posible generalización de determinadas estrategias de aprendizaje por parte de los sujetos, puestas en práctica en su ejecución de las actividades propuestas.

10.4 Discusión de los resultados en las pruebas de la batería criterio

Discusión de los resultados en la prueba de comprensión de textos para educación primaria ECO 1

Con relación al criterio de designación previa del profesorado sobre las dificultades de aprendizaje y, por tanto, teniendo en cuenta el grupo completo de sujetos que formaron parte de los grupos experimental y control del estudio, debemos aceptar la ausencia de incremento significativo en el grupo de tratamiento entre las aplicaciones pretest y postest de esta prueba, y, por el contrario, admitir que el grupo control sí obtuvo mejoras significativas entre las dos fases consideradas.

Como ya quedó expuesto cuando abordábamos el análisis de la equivalencia entre los grupos, tanto el GE como el GC1, es decir, aquel formado por el alumnado con dificultades de aprendizaje, estaban integrados por alumnos y alumnas que formaban, dentro de estos grupos, pequeños pero relevantes subgrupos en función del grado de sus dificultades o de circunstancias relacionadas directa o indirectamente con esta consideración. De este modo, dentro del GE había alumnos que fueron catalogados por sus profesores, o por los orientadores colaboradores, como sujetos que presentaban dificultades de aprendizaje relacionadas con la lectura y la comprensión, en ausencia de cualquier otra circunstancia problemática; otro grupo de alumnos presentaba privación socio-cultural y un tercer grupo de sujetos presentaba algún tipo de discapacidad. Por otra parte, en algunos de los contrastes que realizábamos, habíamos considerado también la circunstancia de recibir o no apoyo a la integración como una variable relevante. Sin embargo, en el GC1, además de los tres subgrupos de alumnos mencionados, hubo que considerar que un cuarto subgrupo lo conformaban aquellos que fueron catalogados por sus profesores tutores como alumnado con dificultades de aprendizaje, pero que no estaban incluidos en ningún tipo de sistemas de apoyo dentro del centro. Esta consideración es importante, por cuanto apuntábamos que las diferencias de nivel de comprensión entre alumnos que asisten a las aulas de apoyo a la integración y aquellos que no lo hacen, podría ser notable. En este sentido, podíamos observar cuando realizábamos los contrastes para valorar la equivalencia de nivel escolar en función de las necesidades educativas especiales relacionadas con los procesos de lectura y comprensión, que el número de sujetos que recibía apoyo a la integración (AI) en el GE era significativamente superior al número de sujetos que recibía este tipo de apoyo en el GC1. Decíamos así mismo que este hecho podría suponer de forma indirecta una falta de equivalencia entre los grupos, lo que se tornaba especialmente relevante en educación primaria, en la que las diferencias encontradas con relación al número de efectivos, eran aún más significativas. En este sentido, a pesar de la no significatividad estadística de las diferencias entre las puntuaciones pretest del grupo de sujetos que recibía apoyo a la integración y las obtenidas por aquellos que recibían apoyo a ciclo o refuerzo educativo, puede deducirse razonablemente en función de los contrastes realizados, que la mayor presencia significativa de alumnado que recibía apoyo a la integración en el grupo experimental, se habría traducido en desventaja con respecto al GC1, y que la comparación con la totalidad de efectivos de la muestra que responde al primer criterio de análisis, podría sufrir un cierto sesgo. La significación de las diferencias pretest-postest, obtenidas por el GC1, podría ser indicativa precisamente de una descompensación relevante en la selección o en la composición de la muestra de control cuando son tenidos en cuenta todos los sujetos designados con dificultades de aprendizaje por el profesorado.

Por otra parte, los resultados de posteriores contrastes en los que tratamos de igualar ambos grupos, nos mostraron que, tanto la mayor presencia de sujetos que recibía apoyo a la integración, como de sujetos con privación socio-cultural en el GE, podría haber influido de forma significativa en la obtención de resultados más pobres en este grupo cuando la comparación se producía entre los grupos completos experimental y control, es decir, teniendo en cuenta todos sus efectivos. Pudimos constatar que el número de efectivos del GC1 descendía notablemente cuando eran eliminados para el contraste los sujetos que no presentaban discapacidad o privación social, o que así mismo no recibían atención educativa especial en las aulas de apoyo a la integración. En cambio, no ocurría de esta forma para el grupo de sujetos del GE, que sólo se veía reducido en dos. Los resultados no sufrían variación para este grupo; en cambio, el GC1 dejaba de

obtener diferencias estadísticamente significativas, aunque seguía manteniendo diferencias importantes favorables al postest, tanto en la puntuación total como en algunas de las cinco dimensiones contempladas.

Ahondando en las posibles explicaciones de los resultados obtenidos para el grupo completo, hemos observado que los análisis efectuados en función de los diversos subgrupos considerados en la muestra, nos mostraron diferencias significativas a favor del GC1 sólo cuando las comparaciones pretest-postest se realizaban con los grupos que no recibían apoyo a la integración, o con los grupos que recibían apoyo a ciclo o participaban de agrupamientos flexibles. Por el contrario, cuando estos contrastes fueron efectuados para los sujetos con discapacidad o para el grupo que recibía apoyo a la integración, no se hallaron diferencias dignas de mención en el GC1. Así mismo, el análisis diferencial de los resultados en función del sexo de los sujetos, nos mostró que las chicas del GC1 obtuvieron ganancias significativas en la prueba, siendo su número notablemente superior al del GE; y que, dentro de este grupo de chicas en el GC1, fue también el subgrupo que no recibía apoyo a la integración, el que explicaría las notables diferencias observadas con el grupo de tratamiento. En cualquier caso, por lo que respecta al grupo experimental, los resultados de los anteriores contrastes son parecidos a los obtenidos para el grupo completo, es decir, se observan pequeños incrementos en algunos de los subgrupos contemplados, pero sin acercarse a la significatividad estadística. Por otra parte, otro elemento del que ya hemos hecho mención, y que podría haber contribuido al incremento de las puntuaciones postest en el GC1, habría consistido en la mayor presencia de sujetos que sólo fueron catalogados por sus profesores con dificultades de aprendizaje, en ausencia de discapacidad o privación socio-cultural.

Una de las conclusiones más relevantes que pueden extraerse de los análisis efectuados para los distintos subgrupos contemplados, es que aquellos sujetos que, en función del criterio de designación del profesorado, presentaban mayores dificultades; básicamente, aquellos que recibían apoyo a la integración y que presentaban discapacidad, no han experimentado incremento en sus puntuaciones postest, ni en el GE ni en el GC1; sin embargo, aquellos que fueron catalogados por sus profesores como alumnos con dificultades de aprendizaje, excluyendo discapacidad o privación socio-cultural, y también aquellos con privación que recibían mayoritariamente refuerzo educativo o participaban de sistemas de apoyo como los agrupamientos flexibles, han sido los que han experimentado una subida significativa o cercana a la significatividad estadística. La literatura informa sobre las distintas necesidades educativas de estos grupos de alumnos y sobre su distinta respuesta a programas o propuestas de intervención educativa (Mora, 1998; Romero, 1999), lo que parece haberse visto reflejado, al menos en parte, en nuestro estudio. De esta forma, este segundo grupo de alumnos al que hemos aludido, habría respondido a un cierto perfil de alumnado que se encuentra ciertamente por debajo del nivel medio del grupo-clase, pero que no presenta grandes dificultades específicas, respondiendo positivamente a una intervención educativa que contemple algún sistema de apoyo o refuerzo educativo. Por el contrario, para el grupo de alumnos que integraba mayormente el grupo experimental, es decir, alumnos con dificultades de aprendizaje que reciben atención educativa específica en aulas de apoyo a la integración y alumnos con discapacidad, se han descrito mayores problemas en la consecución de mejoras cognitivas o académicas mensurables objetivamente. Nuestros resultados, en función de la designación por parte del profesorado de las dificultades de aprendizaje del alumnado, revelan y confirman estos problemas, para los que el

tratamiento dispuesto y aplicado al grupo de tratamiento no ha ejercido la influencia positiva que hubiera sido deseable.

Por otra parte, tal y como suponíamos, los contrastes efectuados en función de los grupos de aplicación del dispositivo, nos muestran la existencia de notables diferencias en cuanto a los resultados obtenidos. De este modo, el factor *aplicador* habría tenido una cierta influencia en las puntuaciones globales del grupo experimental para la prueba de comprensión ECO 1. El análisis de estas diferencias revela que dos grupos experimentan importantes pérdidas, mientras que otros dos obtienen ganancias cercanas a la significatividad estadística en la práctica totalidad de las dimensiones contempladas en la prueba. Así mismo, a efectos exploratorios, procedimos a realizar un contraste de las puntuaciones pretest-postest prescindiendo de los grupos de aplicación en los que habían tenido lugar de forma especial circunstancias que habían influido negativamente, bien en la aplicación misma del dispositivo, o bien en la aplicación de las pruebas criterio. Esta comparación tenía sentido fundamentalmente en función de la determinación de aquellas condiciones, relativas a los sujetos o a los grupos de aplicación, en las que los resultados se tornaran más positivos. El contraste fue efectuado con tres de los grupos de aplicación participantes, otorgando resultados favorables a la hipótesis experimental, y cercanos a la significatividad estadística, con una ganancia media de cuatro puntos.

En función del segundo de los criterios adoptados para el análisis de las diferencias pretest-postest, es decir, la conjunción del criterio de designación previa del profesorado, junto a la puntuación inicial obtenida en la prueba ECO 1, hallamos diferencias significativas favorables para el grupo experimental. Estas diferencias se encontraron para la puntuación total, así como para la dimensión *Metaconocimiento estratégico*. Así mismo, los contrastes mostraron puntuaciones postratamiento claramente superiores a las pretest en casi todas las dimensiones contempladas. Por tanto, podemos deducir de estos datos que la aplicación del dispositivo de evaluación dinámica EDPL mostró un efecto significativo a favor de la primera de nuestras hipótesis experimentales. Estos resultados favorables habrían tenido lugar precisamente cuando la aplicación se había dirigido a los sujetos con mayores dificultades de lectura y comprensión. Por otra parte, la constatación de diferencias muy significativas a favor del postest para el grupo control, tanto en la puntuación total como en tres de las dimensiones, introduce un matiz que ensombrece los datos anteriores. De este modo, el grupo de sujetos que partía de puntuaciones iniciales que se situaban por debajo de una desviación típica de la media, tanto en el grupo experimental como en el grupo de control, obtuvo ganancias significativas en el postest. Estas diferencias se mantenían incluso cuando fueron eliminados de la muestra aquellos sujetos que no recibían apoyo a la integración. Por otra parte, los contrastes efectuados para los compañeros que habían obtenido puntuaciones iniciales más altas, es decir, aquellos sujetos cuyas puntuaciones en el pretest se situaba por encima de los 84 puntos, mostraron pérdidas o ausencia de diferencias significativas. Las razones que hemos esgrimido apuntan al grado de dificultad de la prueba de comprensión con relación al posible margen de ganancia de este subgrupo de sujetos en la fase experimental, unida a la falta de eficacia que habría demostrado el dispositivo EDPL para este grupo de sujetos con mejores puntuaciones de partida.

Por su parte, el contraste efectuado en función del tercero de los criterios considerados, es decir, la designación previa del profesorado más la obtención de una baja puntuación

inicial en comprensión, más una puntuación baja inicial en la prueba de CI, nos reveló que el grupo experimental obtenía ganancias significativas en el postest. Estas ganancias se producían en la puntuación global y en la dimensión *Metaconocimiento estratégico*, siendo también notables en la práctica totalidad de las dimensiones contempladas. En esta ocasión no se encontraron diferencias estadísticamente significativas para el grupo control, a pesar de que las ganancias en el postest fueron importantes. Las razones de esto habría que buscarlas probablemente en el hecho de que haya un sujeto menos en este grupo y en que la variabilidad de sus puntuaciones fuera mayor.

Las puntuaciones de los sujetos que respondían a un perfil caracterizado por presentar dificultades de aprendizaje de la lectura, especialmente relacionadas con los procesos de decodificación o el reconocimiento de palabras; dificultades certificadas tanto por la designación previa del profesorado, como por la obtención de puntuaciones iniciales en la prueba ECO 1 que se situaban por debajo de una desviación típica de la media; y que, en cambio, no presentaron bajas puntuaciones en la prueba de CI (> 84), fueron sometidas a un análisis para comprobar en qué medida estos sujetos se habían beneficiado del tratamiento que suponía su participación de las sesiones de evaluación dinámica. Para este grupo, habitualmente denominado *disléxicos disfonéticos o fonológicos*, compuesto tan solo por cuatro sujetos, se obtuvo una ganancia promedio de 5 puntos en la puntuación global postest, aunque esta diferencia no llegó a ser significativa estadísticamente.

Con relación al análisis de carácter exploratorio efectuado con las DDP en la prueba de comprensión ECO 1; análisis realizado con base en el primero de los criterios formulados, pudimos constatar la ausencia de diferencias dignas de mención entre las puntuaciones de ganancias y de pérdidas. Aunque el contraste con la prueba t muestra una diferencia de medias promedio superior para las ganancias, la escasa magnitud de esta diferencia no añade en este caso información relevante a la obtenida con los contrastes más psicométricos.

Discusión de los resultados en la prueba de comprensión de textos para educación secundaria ECOS

Tomando como criterio la designación del profesorado sobre las dificultades de aprendizaje, pudimos observar diferencias muy significativas a favor del grupo de tratamiento, tanto para la puntuación global de la prueba como para la práctica totalidad de las dimensiones contempladas. Por otra parte, pudimos constatar que la comparación entre las puntuaciones de mejora obtenidas por los grupos experimental y control, mostraba diferencias notables a favor del grupo experimental, llegando a ser estadísticamente significativas para la dimensión *Supervisión y autoevaluación del proceso de comprensión*. De este modo, hemos podido comprobar que era precisamente una de las dimensiones marcadamente metacognitivas la que establecía diferencias nítidas, no solo con relación a las puntuaciones inicial y final del grupo de tratamiento, sino entre las mejoras observadas en los sujetos, en función de su participación o no de las sesiones de aplicación del dispositivo.

Pudimos comprobar la existencia de diferencias significativas o muy significativas, para la puntuación global y para algunas de las dimensiones de la prueba, en los contrastes realizados en función de los distintos subgrupos contemplados. Estas diferencias eran mayores en el grupo experimental, a pesar de lo cual el contraste efectuado para evaluar

la posible significatividad de las diferencias de medias observadas entre el grupo experimental y el de control, ofrece resultados similares a los expresados para el grupo completo de sujetos, es decir, se constata que el grupo de tratamiento obtuvo ganancias de mayor intensidad que el grupo control, aunque sin llegar a ser significativas las diferencias entre ambos grupos.

En función del sexo, los resultados de los contrastes pretest-postest efectuados, mostraron diferencias significativas en todas las comparaciones. Las diferencias más importantes tuvieron lugar entre las puntuaciones pretest y postest de los niños. En este sentido, con relación a la mayor presencia de niños en el grupo control, lo que podía llegar a suponer una desventaja según algunas de las hipótesis que barajábamos en los análisis de nuestra muestra experimental, podemos afirmar que no parece haber ocasionado ningún sesgo real, por cuanto los resultados de los chicos reflejan una mayor diferencia pretest-postest que la obtenida por las chicas.

En cuanto al contraste efectuado en función de los distintos grupos de aplicación, pudimos constatar que las diferencias pretest-postest observadas para el grupo completo de sujetos con dificultades de aprendizaje, fueron refrendadas por los resultados obtenidos en los grupos de aplicación. Es oportuno en este caso recordar que a pesar del reducido número de efectivos en algunos de estos grupos, las ganancias llegaron a ser significativas o muy significativas para la mayoría de ellos, tanto en la puntuación global de la prueba, como en algunas de las dimensiones contempladas. Entre estas dimensiones que obtuvieron ganancias estadísticamente significativas, hay que destacar las de *Conocimientos previos*, *Integración texto-conocimientos* y *Supervisión y Autoevaluación*. De este modo, las mejoras se produjeron precisamente en aquellas dimensiones que habían representado una mayor dificultad para los sujetos. Por otra parte, a través de los análisis efectuados con el fin de comprobar la posible significatividad de las diferencias obtenidas entre los distintos grupos de aplicación, hemos podido constatar que el factor *aplicador* puede llegar a ser un aspecto clave en la explicación de los resultados de la aplicación de un dispositivo de evaluación dinámica como el propuesto. En este sentido, podría constituir un elemento de gran relevancia a la hora de la posible generalización de los resultados, así como de futuras propuestas de formación del profesorado.

Los contrastes realizados en función del segundo de los criterios adoptados, nos mostraron diferencias estadísticamente muy significativas para el grupo de tratamiento, tanto en la puntuación global como en tres de las dimensiones de la prueba. El grupo control, por su parte, también obtuvo ganancias significativas en la puntuación global y en una de las dimensiones, aunque de menor intensidad en ambos casos. Una vez analizada la composición de los grupos experimental y control, se efectuaron nuevos contrastes que excluían a los sujetos que no precisaban de una atención educativa específica, a pesar de haber obtenido puntuaciones iniciales bajas. En este sentido, hemos apuntado a lo largo de este trabajo que esta circunstancia podría ir acompañada de mejores resultados. Por otra parte, el hecho de que algunos alumnos que no requerían este tipo de apoyo, tanto en primaria como en secundaria, hayan obtenido puntuaciones tan bajas en la prueba, podría responder a las dificultades iniciales que la realización de la prueba parece haber tenido para algunos alumnos, dificultades que habrían sido resueltas, a tenor de las puntuaciones postest conseguidas, de forma muy distinta en una y otra etapa. En efecto, tal y como comprobábamos en la submuestra de primaria, estos alumnos a los que hacemos referencia, consiguieron puntuar significativamente mejor

en el postest, resultado que no consiguieron los sujetos que recibieron AI. De esta forma, aquellos alumnos con menores dificultades habían obtenido diferencias mayores que aquellos que, habiendo obtenido también bajas puntuaciones iniciales, mostraban mayores dificultades; dificultades que, como expresábamos anteriormente, se concretaban en su participación de condiciones de apoyo más específicas. Sin embargo, para secundaria no hemos observado esta circunstancia, ya que, a pesar de que las puntuaciones iniciales resultaron ser superiores estadísticamente para el grupo de sujetos que no recibía apoyo a la integración, las ganancias obtenidas por los sujetos de uno y otro subgrupo resultaron estadísticamente significativas.

Los resultados de los contrastes efectuados para evaluar la diferente magnitud de las ganancias observadas en los grupos experimental y control, para el subgrupo de sujetos que puntuó más alto, nos mostraron diferencias significativas favorables al grupo de tratamiento en la dimensión *Supervisión y autoevaluación del proceso de comprensión*. Este resultado vendría a apoyar la tesis de que el dispositivo aplicado tendría una especial incidencia en aquellas dimensiones especialmente relacionadas con los procesos metacognitivos implicados en la comprensión. En definitiva, tanto para la submuestra conformada por sujetos que obtuvieron puntuaciones bajas en ECOS, como para la que lo fue por aquellos que mostraron puntuaciones iniciales que superaban los 84 puntos en esta prueba, las ganancias postest resultaron ser significativas o muy significativas. En ambos casos, la magnitud de las ganancias fue mayor para el GE de forma general, aunque sólo para una de las dimensiones del subgrupo que obtuvo puntuaciones medias o altas, las diferencias resultaron significativas estadísticamente.

El criterio que aúna la designación previa del profesorado, más la puntuación inicial obtenida en la prueba de comprensión, más la puntuación inicial en Cattell, mostró nuevamente diferencias significativas entre el pretest y el postest para los grupos experimental y control, con relación a la puntuación global, obteniendo el grupo de tratamiento ganancias significativas o muy significativas también en dos de las dimensiones contempladas, una de ellas referida al *Metaconocimiento estratégico*. Estos resultados son positivos con relación a verificar nuestra hipótesis principal de trabajo, por cuanto, al igual que ocurriera con los anteriores contrastes realizados en esta etapa, aquellos sujetos a los que fue aplicado el dispositivo de evaluación dinámica EDPL, experimentaron ganancias significativas en sus puntuaciones postratamiento. Los resultados también positivos del GC1 nos hablarían de posibles sesgos en la selección de los sujetos que integraron este grupo, lo que ya analizamos anteriormente en función de los otros criterios seguidos. Por otra parte, la eliminación en los contrastes de los sujetos que no recibían atención educativa específica en las aulas de apoyo a la integración o audición y lenguaje, supuso en este caso, que el grupo de control dejara de mostrar diferencias estadísticamente significativas.

En cuanto a los sujetos que obtuvieron bajas puntuaciones en comprensión, pero alcanzaron puntuaciones medias o altas en CI, los contrastes mostraron ganancias muy significativas en el grupo experimental, tanto para la puntuación global como para algunas de las dimensiones contempladas, entre las que se encontraba nuevamente la de *Metaconocimiento estratégico*. Las ganancias fueron generalizadas en todas las dimensiones. Por su parte, el grupo control, muy reducido, no mostró diferencias pretest-postest dignas de mención.

Por último, los análisis efectuados con el instrumento de las DDP nos permitieron complementar la información obtenida con aquellos más psicométricos. De esta forma, a tenor de los resultados de este análisis, uno de los aspectos más relevantes fue que los efectos positivos de la aplicación dinámica se habrían distribuido entre los efectivos de la muestra, con más del 90% de la evolución de los puntos a favor de nuestra hipótesis experimental, sin que los resultados se debieran a ganancias concentradas en unos pocos sujetos. Las diferencias favorables al grupo de sujetos que experimentó mejoras postratamiento, fueron significativas estadísticamente. Como era de esperar, en función del conjunto de análisis efectuados anteriormente, estos resultados se dieron también para el grupo de control, rozando en este caso la significatividad estadística el contraste de las diferencias promedio entre el grupo de ganancias y el de pérdidas.

Discusión de los resultados en la escala de ajuste personal-social APSL

Tal y como apuntábamos en el capítulo de Resultados, la aplicación de la escala de ajuste personal-social, como forma de medición de los posibles cambios en estos aspectos, no nos permitía efectuar un análisis del proceso de cambio de actitudes, motivación o expectativas en el alumnado. Sin embargo, sí podía ofrecernos una medida del conocimiento y la percepción que los sujetos tenían sobre sí mismos y sobre sus propias actitudes, motivación, intereses, expectativas o modo de afrontar las situaciones problemáticas relacionadas con el aprendizaje de la lectura. De este modo, la escala APSL ha sido considerada un instrumento de evaluación del conocimiento metacognitivo con relación al desarrollo socio-personal. En este sentido, ya hicimos mención de la relevancia de la metacognición como aspecto clave en el desarrollo social-personal y emocional del sujeto; aspecto que podría contribuir significativamente a explicar las dificultades de aprendizaje, así como los trastornos relacionados con el ajuste personal-social (Mora, 1994; Mora Merchán y Mora, 2000) en el caso de que su funcionamiento presente problemas.

Las puntuaciones obtenidas por los sujetos en la escala APSL debían ofrecernos esta información sobre el conocimiento y la percepción que muestran los sujetos acerca de sus propias dificultades, su motivación, sus intereses o sus expectativas. Así mismo, las comparaciones entre los puntajes iniciales y los de postratamiento, nos otorgarían información acerca del posible grado de mejora observado en el grupo de tratamiento, en función de su participación en las sesiones de aplicación del dispositivo, y con relación al cambio observado en ausencia de tratamiento. La aplicación del dispositivo debía originar cambios a favor de la hipótesis experimental, aún cuando estos no fueran muy significativos, debido a los condicionantes que ya expusimos en el citado capítulo de Resultados, así como al propio contenido objeto del análisis. En este sentido, en función del criterio de designación previa del profesorado sobre las dificultades de aprendizaje, debemos reseñar la ausencia de ganancias significativas para el grupo experimental, aunque son destacables las mejoras postratamiento en la puntuación global, y de forma especial en algunas de las dimensiones contempladas, entre ellas la dimensión D9 (*Reconocimiento de la capacidad de generalización de conocimientos y estrategias*). Por su parte, el grupo de control experimentó pérdidas postest, tanto en la puntuación global como en la mayoría de las dimensiones de la escala.

Los análisis efectuados con las DDP nos mostraron diferencias aún más nítidas entre los resultados obtenidos por los grupos experimental y control. De este modo, el grupo de tratamiento obtenía pequeñas diferencias promedio entre el pretest y el postest, tanto en

los análisis psicométricos como en los contrastes con las DDP, mientras que el grupo control obtenía pérdidas, que en el caso de los análisis con la prueba exploratoria, mostraba diferencias porcentuales de 20 puntos en cuanto al número de sujetos para el que decrecían sus puntuaciones posttest, y de 16 puntos cuando se tenía en cuenta la magnitud de las ganancias o de las pérdidas en el análisis.

Por otra parte, los contrastes efectuados en función de la etapa nos han permitido profundizar en el análisis del comportamiento diferencial de los sujetos de primaria y secundaria. La previsión de la existencia de claras diferencias en cuanto al nivel de ajuste mostrado en la prueba, así como en el modo de afrontamiento de la actividad propuesta, dada su naturaleza metacognitiva (reflexión sobre sus propias actitudes, su motivación, sus intereses o expectativas, su percepción del apoyo que recibe, etc.), se vio confirmada con algunos de los resultados obtenidos. En este sentido, pudimos constatar diferencias estadísticamente significativas en el alumnado de secundaria para algunas de las dimensiones contempladas en la escala. Las ganancias observadas en los sujetos mayores son especialmente relevantes en las dimensiones relacionadas con el *metaconocimiento estratégico sobre la lectura* (D8) y con el *reconocimiento de la capacidad de generalizar esas estrategias* (D9), llegando a ser muy significativas para ésta última, así como para la dimensión que aúna los aspectos citados (D8 y 9). Como ya apuntábamos en nuestra exposición de los resultados, estos resultados nos permiten matizar los resultados globales y nos ayudan a delimitar mejor el perfil de los sujetos que podrían haberse beneficiado en mayor medida de la aplicación del dispositivo y, especialmente, del trabajo desarrollado con relación a los procesos de ajuste personal-social. Por otra parte, estos resultados nos hablan también de las dimensiones sobre las que la actuación desarrollada habría tenido un mayor efecto. De este modo, volvemos a obtener datos que apuntan hacia la relevancia que los procesos metacognitivos tienen en el dispositivo de evaluación dinámica desarrollado y puesto en práctica, así como hacia la relevancia que tendrían los procesos de ajuste en el rendimiento observado en lectura. Las mejoras observadas en las dimensiones metacognitivas de la escala, se han producido precisamente en la etapa donde tuvieron lugar las mayores ganancias postratamiento en comprensión.

Hacíamos alusión cuando analizábamos la equivalencia entre los grupos experimentales, a que un número importante de estudios avalaban mejores niveles de adaptación social y escolar para las chicas, lo que podía traducirse en actitudes más positivas en el contexto escolar, así como en un grado de motivación más óptimo con relación al trabajo desarrollado en la escuela (Freixas y Luque, 1998; del Río, 1999; Mora, 1998). En este sentido, tomando la muestra completa de sujetos, es decir, el grupo experimental más el grupo control, no pudimos encontrar diferencias significativas entre chicos y chicas; tampoco hallamos diferencias dignas de mención cuando se efectuaron los contrastes en función de la etapa educativa. De este modo, los resultados obtenidos en la escala parecen contradecir inicialmente lo apuntado en referencia al mayor rendimiento en lectura, y también en cuanto a la mejor adaptación social y escolar de las niñas, en la etapa de primaria. En este sentido, es importante destacar que las investigaciones citadas con relación a esta supremacía de las chicas en el ámbito que nos ocupa, están fundamentadas básicamente en muestras de población general, es decir, sin un sesgo previo en función precisamente de presentar dificultades de aprendizaje de la lectura. Sin olvidar que en las comparaciones realizadas por nosotros para la etapa de primaria, el número de niñas fue sensiblemente menor que el de niños, y que en cuanto a la puntuación global no se obtuvieron diferencias significativas entre chicos y chicas para

ninguno de los contrastes efectuados, nuestros resultados apuntarían hacia un igualamiento entre chicos y chicas en rendimiento en lectura y adaptación personal-social cuando existen dificultades de aprendizaje relacionadas con la lectura. Esta igualdad puede llegar a tornarse en puntuaciones superiores en los chicos cuando se trata de las dimensiones que hacen referencia al *conocimiento de las propias expectativas, autoconcepto académico y estilo atribucional* (D2), así como al *conocimiento de la propia actitud ante la lectura: emociones, sentimiento de rechazo versus satisfacción y gusto por la lectura, tolerancia a la frustración, persistencia ante la dificultad, autonomía versus dependencia* (D3).

Apuntábamos también que el mejor desempeño de las chicas en cuanto al ámbito del lenguaje escrito podía reducirse y llegar a igualarse con los chicos al término de la etapa de secundaria. Si bien esta tendencia es observable de forma general entre la población adolescente, aspecto que de alguna forma también recogen nuestros resultados con relación a las comparaciones efectuadas con la puntuación global, es necesario destacar que las chicas con dificultades de aprendizaje han obtenido puntuaciones superiores en la etapa de secundaria. Este hecho no nos sorprende del todo, por cuanto se ha descrito en numerosas ocasiones que los niveles de adaptación socio-personal y escolar de los chicos, especialmente en los primeros cursos de esta etapa, estarían sensiblemente por debajo de los niveles que presentan habitualmente las chicas, mostrando de forma significativa un mayor número de desajustes con el contexto escolar y social, desajustes que se concretan a menudo en un mayor número de conductas disruptivas, unos mayores índices de absentismo escolar o un mayor número de enfrentamientos con el profesorado, la familia y el centro escolar. Este dato podría ser incluso más relevante cuando se hace referencia al alumnado que presenta dificultades de aprendizaje. En este sentido, las chicas habrían mostrado a nivel general, e incluso de forma significativa en alguna de las dimensiones contempladas en la escala, un mayor nivel de ajuste personal-social medido con APSL. En línea con lo expuesto anteriormente, la dimensión en la cual se han obtenido estas diferencias estadísticamente significativas con respecto a la puntuación de los chicos, no parece en absoluto casual; se trata de la D5, que hace referencia a la *percepción del apoyo que ofrece el contexto escolar y familiar, y de las condiciones en las que se desenvuelve*.

El análisis de los contrastes efectuados en función de los distintos subgrupos considerados en la muestra, ha servido para ampliar de forma significativa el alcance de nuestras conclusiones. Comentábamos en el capítulo de Resultados que el dispositivo en su conjunto, así como el trabajo realizado con relación a los procesos de ajuste, podrían haber ejercido una influencia diferencial en los sujetos participantes, en función del grado de sus dificultades o de las condiciones de apoyo que tuvieran. En este sentido, pudimos comprobar que para el alumnado que solo presentaba dificultades de aprendizaje, en ausencia de discapacidad o deprivación, los resultados mostraban ganancias postratamiento para el grupo experimental en la puntuación global, mientras que indicaban pérdidas para el grupo de control. Cuando se analizaban las distintas dimensiones de la escala, constatábamos de nuevo que eran las dimensiones metacognitivas las que obtenían diferencias más relevantes entre pretest y posttest para el grupo de tratamiento, obteniéndose diferencias estadísticamente significativas para la dimensión D8 y 9. También pudimos comprobar que las puntuaciones de mejora establecían diferencias importantes entre los grupos experimental y control, lo que sin duda es un dato de especial relevancia. En este sentido, el contraste efectuado mostraba que, tanto en la puntuación global como en la mayoría de las dimensiones, las

diferencias promedio de las mejoras eran favorables al grupo experimental, llegando a ser significativas para la dimensión D1 (*Conocimiento de las propias limitaciones cognitivas actuales, potencialidades y posibilidades de aprendizaje, dificultades y vías de solución*). Sin embargo, los contrastes efectuados para el segundo de los subgrupos, compuesto por los sujetos con discapacidad, arrojó resultados menos alentadores. Los datos mostraron pérdidas postest para el grupo experimental, tanto en la puntuación global como en las dimensiones de la escala. Por su parte, el grupo control también obtuvo pérdidas en este contraste.

Con relación al subgrupo formado por el alumnado con deprivación socio-cultural, los resultados mostraron pequeñas ganancias postratamiento para el grupo experimental, aunque éstas no llegaron a ser significativas en ninguna de las puntuaciones examinadas. El grupo control tuvo un comportamiento parecido, con pequeñas ganancias postest, que en ningún caso se acercaban a la significatividad estadística. Los contrastes efectuados en función de la condición *recibe apoyo a la integración*, nos mostraron igualmente ausencia de diferencias significativas entre las puntuaciones iniciales y postratamiento, tanto para el grupo experimental como para el grupo de control. Sólo los valores de diferencia en la dimensión metacognitiva D9 se acercaron a la significatividad estadística. En lo que respecta al grupo de sujetos que no recibía apoyo específico, ni presentaba discapacidad o deprivación socio-cultural, volvimos a constatar el contraste entre la cantidad de efectivos que componían estas submuestras en los grupos experimental y control, a favor del segundo. Este dato podría relacionarse con mayores puntuaciones postest para el grupo control, en virtud de una mayor modificabilidad de los sujetos que no presentan discapacidad y no precisan una atención educativa especializada. A pesar de esto, los datos nos mostraron ganancias para el grupo experimental, aunque éstas no llegaban a ser significativas; y, por el contrario, indicaban pérdidas para el grupo de control. La mayoría de las dimensiones obtenían mejoras por encima de los 8 puntos, siendo nuevamente las dimensiones metacognitivas D8 y D9 las que obtuvieron diferencias pretest-postest más cercanas a la significatividad. Estos datos quedan reforzados por cuanto el grupo de control obtuvo pérdidas, aunque no significativas, en las distintas puntuaciones, así como por el hecho de que el contraste entre las puntuaciones de mejora de ambos grupos arrojó resultados muy favorables al grupo de tratamiento, con diferencias promedio muy pronunciadas para algunas de las dimensiones, entre las que también se encontraba la D9.

Por otra parte, la realización de contrastes en función de los grupos de aplicación del dispositivo, suponía la posibilidad de analizar los resultados obtenidos por cada grupo de sujetos a los que un mismo profesional aplicaba EDPL. Así mismo, otro de los motivos relevantes de estos análisis consistía en contrastar los resultados obtenidos en las distintas pruebas criterio por un mismo grupo. Estos análisis nos proporcionarían datos importantes acerca de la validez de las ganancias observadas. En este sentido, a pesar de que para esta prueba no se pudieron constatar diferencias estadísticamente significativas favorables al grupo de tratamiento, sí se observaron notables incrementos postest para la mayoría de los grupos de aplicación en algunas de las dimensiones de la escala. Las mayores diferencias, así como el mayor número de ganancias en los distintos grupos, se dieron en las dimensiones metacognitivas: D8, D9 y D8 y 9, así como también en la D1.

El hecho de que sea en estas dimensiones de metaconocimiento estratégico sobre la lectura, en las que un mayor número de grupos de aplicación haya obtenido ganancias

postest —algunas de éstas con notables diferencias de promedio—, nos conduce a una doble reflexión. Por una parte, estas dimensiones hacían referencia, como apuntábamos cuando analizábamos la escala APSL, a contenidos que también podíamos encontrar en las pruebas de comprensión de textos, ya que en éstas, había también dimensiones relativas al metaconocimiento de estrategias de comprensión; es decir, se trataría de contenidos más cercanos a lo conceptual —en este caso, a lo metaconceptual— que al conocimiento sobre las propias actitudes, motivaciones e intereses. Una posible interpretación de estos resultados podría tener que ver con el hecho de que ha sido precisamente sobre los contenidos relacionados con los metaconocimientos sobre estrategias, sobre los que ha tenido una mayor repercusión el trabajo realizado mediante la aplicación del dispositivo de evaluación dinámica. Por otra parte, habíamos apuntado también las posibles dificultades que podrían encontrarse para observar cambios relevantes en procesos de ajuste más relacionados con las actitudes y la motivación. De esta forma, otra posible explicación, que podría ser complementaria a la anterior, pondría el énfasis en que los procesos de ajuste habrían sufrido pocos cambios cuantificables en el tiempo transcurrido durante la aplicación experimental, aunque, como veremos más adelante, la valoración cualitativa de los aplicadores y la evaluación de la aplicación del dispositivo, nos pudieran ofrecer datos para la controversia; en cambio, los procesos metacognitivos, evaluados también mediante las dimensiones D8, D9 y D8 y 9 de la escala APSL, más susceptibles de modificación a corto y medio plazo, sí habrían visto transformados positivamente sus valores postratamiento.

Los análisis efectuados en función del segundo de los criterios adoptados, mostraron diferencias muy significativas para el grupo experimental con relación a la puntuación global de la escala, así como para algunas de las dimensiones contempladas, entre las que vuelven a estar la D1 y la D8 y 9. Por su parte, el grupo de control no obtuvo diferencias estadísticamente significativas para ninguna de las puntuaciones analizadas. El contraste entre las mejoras obtenidas por ambos grupos no mostró significatividad estadística, a pesar de que podían observarse notables diferencias entre sus diferencias promedio pretest y postest, tanto para la puntuación global, como para la mayoría de las dimensiones. Por otra parte, el análisis de regresión nos permitió comprobar que las puntuaciones iniciales obtenidas en la escala, explicaban de forma significativa estos resultados positivos obtenidos por el grupo experimental. Al mismo tiempo, el porcentaje de varianza no explicada por el pretest nos hablaba de la posible efectividad del tratamiento aplicado, y de la información adicional y complementaria que éste aportaba a la obtenida con la pasación de la escala.

Los contrastes efectuados para aquellos sujetos que obtuvieron bajas puntuaciones en comprensión, arrojaron resultados favorables al grupo experimental en primaria, con una puntuación de mejora que superaba los dos puntos y medio para la puntuación global en la escala, y que era superior para algunas de las dimensiones contempladas. El grupo de control no obtuvo incremento postest. Por su parte, el grupo experimental de secundaria obtuvo una mejora cercana a los 5 puntos, aunque la diferencia no llegó a ser significativa, debido probablemente a lo exiguo de la muestra. Sí se encontraron diferencias significativas para la dimensión D3. En cuanto al grupo de control, éste obtuvo en esta ocasión pérdidas notables en el postest. Por otra parte, los contrastes efectuados para los sujetos que obtuvieron bajas puntuaciones, tanto en la prueba de comprensión como en la escala de ajuste, mostraron grandes diferencias pretest-postest a favor del grupo experimental. Estas ganancias, a las que hemos hecho referencia, obtenidas por el grupo de tratamiento en la escala APSL, se produjeron en primaria y en

secundaria, aunque lo reducido de la muestra que reunía las condiciones de estas comparaciones, no permitía la significatividad estadística de las diferencias observadas. Así mismo, estos resultados no avalaban la extracción de conclusiones potentes y generalizables, más allá de la hipótesis de la posible efectividad del dispositivo para contribuir al incremento de algunas de las dimensiones del proceso de ajuste personal-social, precisamente en aquellos sujetos que presentaron puntuaciones especialmente bajas en comprensión, y también en aquellos que además también puntuaron inicialmente bajo en la propia escala.

En función del criterio que aúna la designación previa del profesorado, junto a las puntuaciones iniciales obtenidas en la escala y en la prueba de CI, los resultados mostraron nuevamente ganancias significativas a favor del grupo experimental para la puntuación global y para las dimensiones D1, D9 y D8 y 9. El grupo control, por su parte, obtuvo también importantes diferencias pretest-postest, aunque éstas solo llegaron a ser estadísticamente significativas para la dimensión D2. El análisis de regresión mostraba que las puntuaciones iniciales de la escala y de la prueba de CI explicaban de forma conjunta el 67% de la puntuación postest, aunque el modelo resultante no lograba explicar significativamente los resultados obtenidos en el postest. Pudimos comprobar así mismo, que las puntuaciones iniciales en CI ocasionaban, con su introducción en la ecuación, un decremento en el porcentaje de varianza explicada por el modelo. Una conclusión que podemos extraer de estos análisis es relativa a que el nivel de inteligencia en el pretest, no predice bien la puntuación postest obtenida en la escala APSL, y contribuye a desestimar el modelo cuando se introduce junto a la variable relativa a la puntuación inicial en la escala. Otra conclusión hace referencia a la insuficiencia de las puntuaciones iniciales en APSL y en Cattell para explicar completamente los resultados obtenidos tras la aplicación del dispositivo EDPL.

Con relación a los análisis efectuados para evaluar la evolución pretest-postest de las correlaciones entre las puntuaciones obtenidas en la escala de ajuste y en las pruebas de comprensión, pudimos constatar que para la prueba de comprensión ECO 1, la evolución de las correlaciones fue positiva para el grupo experimental en los tres criterios adoptados, y especialmente para el que aunaba la designación previa del profesorado con la obtención de puntuaciones inicialmente bajas en la escala. Por el contrario, para el grupo de control la evolución de las correlaciones fue marcadamente negativa. En cuanto a la etapa de secundaria, los resultados de estos análisis para el grupo experimental no mostraron una evolución tan positiva de las correlaciones entre ambas pruebas. A pesar de que se produjeron incrementos en las correlaciones para los dos primeros criterios, estos fueron de escasa magnitud, mientras que para el tercer criterio, se produjo un decremento en la correlación. Sin embargo, en esta etapa, el grupo control mostró una evolución más positiva de las correlaciones. El análisis de la evolución de las correlaciones entre la escala de ajuste y las pruebas de comprensión, nos ofrece de este modo información complementaria y valiosa acerca de una evolución positiva en el grado de ajuste personal-social con relación a la lectura para el grupo experimental de primaria a lo largo de la aplicación del dispositivo EDPL. Estos resultados complementan y matizan los obtenidos mediante análisis más psicométricos, evitando posibles sesgos relacionados con una apreciación y valoración de los datos de la escala sin tener en cuenta la evolución de las puntuaciones obtenidas en comprensión. En secundaria, los resultados positivos obtenidos por la mayoría de los sujetos del grupo experimental, tanto en la prueba de comprensión como en la escala de ajuste, no habrían

tenido su reflejo en una evolución más marcadamente positiva de la correlación entre el ajuste y la comprensión.

Por otra parte, en el capítulo de Resultados, hacíamos referencia a la relevancia de la evaluación de la fiabilidad de los cambios observados, en el análisis de los instrumentos de evaluación dinámica (Sternberg y Grigorenko, 2003). Entre los métodos propuestos para valorar este aspecto, nombrábamos el citado por Embretson (1987), que establecía la utilización de tres postest diferentes para un mismo pretest. Destacábamos que la utilización conjunta de las pruebas de ajuste y comprensión podría ofrecer datos relevantes sobre la convergencia de los resultados obtenidos. El contenido de algunas de las dimensiones que comparten ambas pruebas, así como el hecho de que ambas evalúen procesos implicados en la lectura que estaban contemplados en nuestra teoría de base, son aspectos que sustentan la oportunidad, al tiempo que la relevancia de estos análisis.

Los resultados con relación al grupo de primaria, apuntan, como ya señalábamos anteriormente, hacia un mayor ajuste entre las puntuaciones globales obtenidas en ambas pruebas, con posterioridad a la aplicación del dispositivo experimental; así mismo, para la D2 de APSL, la convergencia con la puntuación global de comprensión llega a ser significativa. Por otra parte, con relación a los contrastes efectuados entre las puntuaciones de mejora en APSL y la puntuación total en comprensión, es sumamente destacable que los sujetos que obtuvieron mejoras postratamiento en las dimensiones metacognitivas, obtuvieron puntuaciones más altas en el postest de la prueba de comprensión ECO 1. Este último dato tiene una especial relevancia, por cuanto expresaría convergencia entre las puntuaciones obtenidas en las dimensiones más «conceptuales» de la escala y las obtenidas en la prueba de comprensión, lo que a su vez contribuiría a establecer una mayor fiabilidad de las mejoras observadas. A pesar de estos resultados, las conclusiones generales que podemos efectuar, con base en la totalidad de los análisis realizados, no apuntan hacia índices potentes de fiabilidad del cambio para las pruebas ECO 1 y APSL. Las débiles correlaciones obtenidas en algunos de los diversos contrastes realizados, así como el exiguo número de sujetos que obtuvieron ganancias en ambas pruebas, parecen confirmar que las mejoras observadas a nivel del grupo completo, no son especialmente convergentes cuando se utilizan las dos pruebas para valorarlas.

Los resultados de los análisis efectuados en secundaria, mostraron paradójicamente que, a pesar de la obtención de mejoras por parte del grupo experimental en las puntuaciones medias de la dimensión *Metaconocimiento estratégico* de la prueba de comprensión, así como en las dimensiones metacognitivas de la escala APSL, los cambios positivos observados no se han producido en una misma dirección. El análisis de las correlaciones efectuadas entre las puntuaciones obtenidas en las dimensiones metacognitivas de ambas pruebas, así como entre las puntuaciones de mejora, no ofreció valores especialmente positivos, tal y como podía suponerse a tenor de los buenos resultados de este grupo en ambas pruebas.

Como conclusión de este apartado, y en función de la valoración que hemos realizado de la convergencia de las mejoras obtenidas en las pruebas de comprensión y de ajuste personal-social, podríamos decir que no estamos en condiciones de confirmar la fiabilidad de estas mejoras a tenor de los análisis efectuados. Aunque es cierto que, sobre todo en secundaria, un número importante de sujetos obtiene ganancias en ambas pruebas, dato quizás más relevante que las correlaciones observadas entre las

puntuaciones o entre las diferencias promedio, este dato no parece suficiente. Una posible explicación a la diferente dirección que han tomado las mejoras de los sujetos ganadores, estriba en que las diferencias entre los procesos valorados por ambas pruebas, así como entre los formatos de aplicación, hayan tenido un peso significativo para el desempeño de los sujetos. Este hecho parece haber tenido una relevancia notable precisamente en secundaria, donde a tenor de los resultados de las correlaciones efectuadas, muchos sujetos que mejoraron sus puntuaciones en comprensión no lo hicieron sin embargo en la escala de ajuste personal-social. Esto ocurre, tanto en el grupo completo de sujetos, como en la submuestra compuesta por aquellos que obtuvieron inicialmente bajas puntuaciones en comprensión. Por ejemplo, para el grupo completo, mientras que el 84% de sujetos obtenía ganancias posttest en ECOS, tan solo el 46% de los mismos obtenía mejoras postratamiento en APSL. Esta diferencia se acentúa cuando se adopta el segundo de los criterios; en este caso, fue el 92% de los sujetos que obtuvieron bajas puntuaciones iniciales, el que consiguió mejoras posttest en comprensión, mientras que sólo el 36% de estos alumnos obtuvo ganancias en la escala de ajuste. Otro dato en este sentido, que a su vez confirma los anteriores, es que de los 21 sujetos que obtienen ganancias posttest en la prueba ECOS, el 42% consiguió mejoras en la escala APSL; este porcentaje, que no parece en sí mismo bajo, deja al descubierto sin embargo que el 58% restante obtuvo pérdidas o ganancias nulas, lo que explicaría las correlaciones negativas o escasamente significativas. En el otro sentido, es decir, seleccionando ahora los sujetos que obtuvieron ganancias en APSL, obtenemos que de los 12 alumnos que ganaron en el posttest, cerca del 73% obtuvo también mejoras postratamiento en la prueba ECOS. Así pues, parece ser que las ganancias obtenidas en la prueba de comprensión para la etapa de secundaria, no se han visto acompañadas en magnitud por las mejoras observadas en la escala de ajuste.

Discusión de los resultados en el test de inteligencia general de Cattell

Los resultados obtenidos en los distintos contrastes realizados, con relación a la prueba de inteligencia general de Cattell-2, verifican de forma general y sustancial la tercera de nuestras hipótesis. Ésta hacía referencia a la posibilidad de que los sujetos del grupo experimental vieran incrementado su rendimiento cognitivo en la prueba de CI como consecuencia de su participación en las sesiones de trabajo con el dispositivo de evaluación dinámica EDPL. Pudimos comprobar que para los distintos criterios adoptados, las ganancias postratamiento que experimentaba el grupo experimental completo, fueron muy significativas estadísticamente, oscilando las diferencias promedio entre los 5 puntos, cuando el criterio fue la designación previa del profesorado acerca de las dificultades del alumnado, y los 8 puntos CI cuando el criterio aunaba la designación previa del profesorado, la puntuación inicial en la prueba de comprensión y la puntuación inicial en Cattell. Por su parte, en estas comparaciones, el grupo control no obtenía ganancias significativas, e incluso llegaba a obtener ligeras pérdidas posttest. En cuanto a los contrastes realizados con el grupo reducido, es decir, con aquellos sujetos que recibían apoyo a la integración y presentaban discapacidad o deprivación socio-cultural, las ganancias del grupo experimental fueron aún mayores, mientras que el grupo de control obtenía pérdidas. Las ganancias significativas del grupo experimental se produjeron tanto en la muestra de primaria como en la de secundaria. Los resultados positivos del grupo experimental también se hicieron extensivos a los análisis efectuados en función de los distintos subgrupos contemplados, aunque tan sólo el subgrupo de sujetos con discapacidad obtuvo ganancias muy significativas.

Por otra parte, la constatación de que aquellos sujetos con mejores puntuaciones iniciales, tanto en comprensión como en CI, así como aquellos que no precisaban de apoyo a la integración ni presentaban discapacidad, no obtenían resultados tan satisfactorios, parece confirmar una cierta tendencia relativa a que la aplicación del dispositivo de evaluación dinámica EDPL habría tenido un mayor efecto en la población de sujetos con mayores dificultades y que partían de puntuaciones más bajas. Como ya apuntábamos con anterioridad, la relevancia de estos datos se torna especialmente importante, por cuanto está en consonancia con los resultados observados en muchas de las investigaciones relacionadas con la implementación de procedimientos dinámicos de evaluación. De esta forma, la intervención que supone la aplicación de EDPL, aunque haya contribuido también a la obtención de mejoras en los sujetos con mayores puntuaciones iniciales, no habría tenido para este subgrupo el peso y la significatividad que sí ha tenido con relación a los sujetos que puntuaban bajo. En este sentido, la obtención de mayores índices de mejora en el grupo de control, con relación al grupo de tratamiento, precisamente cuando se analizaron los subgrupos de sujetos que partían de puntuaciones altas (por encima de 84 puntos), tanto en comprensión como en CI, parece apuntar al efecto diferencial del dispositivo en función del grado de dificultad de los sujetos.

Los análisis efectuados con el procedimiento de las DDP ilustraron los resultados obtenidos en los análisis psicométricos, en el sentido de ofrecer datos relativos al porcentaje de sujetos de los grupos experimental y control que obtuvieron las ganancias postratamiento, así como al porcentaje de puntos indicadores de ganancias o de pérdidas. En este sentido, pudimos comprobar que más del 70% de los sujetos del grupo de tratamiento obtuvieron ganancias posttest y que más del 76% de la evolución de los puntos fue favorable a la hipótesis experimental formulada. Por su parte, los resultados para el grupo control muestran ausencia de diferencias relevantes entre los porcentajes de sujetos que obtienen ganancias y pérdidas, así como entre los porcentajes que incorporaban la magnitud de las ganancias y las pérdidas. Con relación al porcentaje de sujetos que obtuvieron pérdidas en el grupo experimental, los análisis anteriormente expuestos apuntan claramente hacia el alumnado que obtuvo puntuaciones iniciales por encima del valor de corte, es decir 84. Hemos anotado, tanto en el capítulo de Resultados como aquí mismo, la posibilidad de que los efectos positivos de la aplicación del dispositivo EDPL con relación al rendimiento cognitivo en la prueba de CI, se hayan producido especialmente en el grupo de sujetos con mayores dificultades y que obtuvieron, a su vez, puntuaciones iniciales más bajas, tanto en comprensión como en Cattell. Los contrastes efectuados con las DDP no hacen sino confirmar estos hallazgos, al tiempo que nos ayuda a detenernos en aquellos sujetos que obtuvieron pérdidas postratamiento, con el consiguiente análisis de las razones que podrían estar detrás de estos resultados, así como de la magnitud de dichas pérdidas.

Apuntábamos anteriormente que en el transcurso del proceso de investigación, contemplamos la posibilidad de que el trabajo sobre procesos metacognitivos, aunque en este caso se tratara de un trabajo encaminado específicamente a la mejora de los procesos de comprensión de textos, pudiera traducirse en la obtención de mejoras en una prueba de factor «G», de las denominadas «libres de cultura» en función de su contenido no verbal. Como recordaremos, cuando abordábamos el análisis de los modelos de evaluación dinámica en nuestro marco teórico, y hacíamos referencia a los estudios de Fernández-Ballesteros y sus colegas (1997), en los que éstos hallaban resultados posttest positivos en pruebas con contenido verbal, tras haber aplicado la fase

de intervención de la EPA, basada en el entrenamiento sobre el test de Raven; hacíamos alusión a que estos resultados, en principio no esperados, suponían una transferencia entre los ámbitos verbal y no verbal, que los autores explicaban con base en el contenido *verbal* de la instrucción en las matrices de Raven, así como apelando a las afirmaciones de Campione y Brown (1979), según las cuales lo que verdaderamente se trabaja cuando se aplican pruebas dinámicas serían procesos de control ejecutivo y regulación verbal.

Al hilo de estas reflexiones, nosotros apuntábamos entonces que la instrucción en estrategias metacognitivas, ya fuera de forma más o menos consciente o intencionada, por parte de los investigadores, podría explicar ciertamente esta transferencia entre ámbitos verbal y no verbal. Así mismo, este hecho podría tener lugar en nuestro estudio, aunque en un sentido de transferencia inverso. En efecto, la fase de aplicación del dispositivo ha actuado fundamentalmente sobre procesos implicados en la lectura y la comprensión, y, por tanto, sobre un ámbito fundamentalmente verbal; sin embargo, el trabajo basado esencialmente sobre los procesos metacognitivos de control y regulación consciente, centrados por tanto más sobre estrategias más generales para la resolución de problemas que sobre funciones cognitivas específicas, podría explicar los posibles cambios en la actuación con pruebas de contenido no verbal. Los sujetos, de esta forma, habrían aprendido estrategias para resolver problemas de diversa índole, porque la esencia de su aprendizaje estaría en la aplicación de elementos generales como, por ejemplo: (a) una mayor reflexión previa a la actuación, examinando el tipo de tarea, comprobando elementos conocidos o conectando el contenido con conocimientos previos; (b) el mantenimiento de la atención por más tiempo y con mayor calidad; (c) ejerciendo un mayor control y regulación sobre la actividad mientras está ejecutándola; y (d) revisando su actividad una vez terminada, y corrigiendo posibles errores. La aplicación de estas estrategias a la realización de una prueba como Cattell, estrategias que habrían sido aprendidas o mejoradas durante el trabajo desarrollado con el dispositivo de evaluación dinámica EDPL, habría contribuido de forma significativa a la explicación de las mejoras obtenidas en el postest de esta prueba.

Discusión de los resultados de la muestra piloto en el postest y de los análisis realizados en función de la Teoría de Respuesta al Ítem

A la luz de la Teoría de Respuesta al Ítem (TRI) (Embretson y Reise, 2000; Hambleton, Swaminathan y Rogers, 1991), resultaba especialmente interesante comprobar en qué tipo de ítems obtenían sus mejoras los diversos subgrupos que podían establecerse, en función de la catalogación como sujetos con dificultades de aprendizaje, así como de las puntuaciones obtenidas en la fase pretest. En este sentido, los análisis efectuados con relación a esta teoría, nos mostraron para el grupo experimental de primaria, que aquellos sujetos que habían obtenido puntuaciones iniciales bajas en la prueba de comprensión y que, a su vez, obtuvieron ganancias significativas postratamiento, mostraron mejoras en todos los ítems que habían sido catalogados como «fáciles» en función de los resultados de la muestra piloto. Así mismo, estos sujetos *mejoradores* obtuvieron ganancias postest en otros ítems catalogados como difíciles o muy difíciles. De esta forma, las ganancias de este subgrupo no se limitaron a los ítems en teoría más fáciles de la prueba, sino que se produjeron también en otros ítems de la misma, en principio difíciles incluso para aquellos sujetos sin dificultades de aprendizaje. El contraste de las puntuaciones pretest y postest para el grupo completo en esta etapa, teniendo en cuenta la valoración de la

dificultad de los ítems, no produjo resultados novedosos con respecto a los ya reflejados, es decir, ausencia de ganancias significativas para el grupo experimental.

En cuanto al grupo de secundaria, los análisis efectuados destacan que los ítems «fáciles» supusieron en su gran mayoría —como era de esperar— la obtención de ganancias en los contrastes pretest-postest para los *mejoradores*. También para este grupo, las ganancias se produjeron, no solo en estos ítems, sino en aquellos catalogados como difíciles o muy difíciles en función de la aplicación piloto realizada. Es más, en este caso, la mayoría de los ítems que otorgaron mejoras postest para este grupo, fueron considerados «difíciles» en función de su promedio de aciertos en la muestra piloto. En este caso, el contraste pretest-postest efectuado para el grupo experimental completo en esta etapa, teniendo en cuenta sólo aquellos ítems catalogados como «fáciles», mostró diferencias significativas favorables al grupo de tratamiento. Por otra parte, dado que este grupo obtenía ganancias muy significativas con el conjunto total de ítems de la prueba; y que, por otra parte, el subgrupo de *mejoradores* había obtenido la mayoría de sus mejoras postest en ítems catalogados como difíciles, podemos decir sin reservas que las mejoras observadas en las pruebas de comprensión se han producido, no sólo en aquellos ítems considerados «fáciles», sino que éstas también tuvieron lugar en ítems que resultaron especialmente difíciles, incluso para los sujetos que no presentaban dificultades.

Con relación al grupo de aplicación piloto, los análisis efectuados mediante la Teoría de Respuesta al Ítem también nos han permitido conocer el tipo de ítems en el que este grupo obtenía sus mejoras. Ya comentamos que, supuestamente, la mayoría de las ganancias debería corresponder a los ítems catalogados como «difíciles», ya que aquellos considerados más fáciles habrían sido resueltos sin problemas en la fase pretest. En este sentido, los análisis realizados nos mostraron que efectivamente los ítems más difíciles de la prueba habían otorgado la mayoría de las mejoras postest para la muestra piloto de primaria. En efecto, más del 81% de los ítems de mejora habían sido catalogados como difíciles o muy difíciles. Por otra parte, las ganancias para este grupo se produjeron de forma generalizada, teniendo lugar también en aquellos ítems de la prueba considerados «fáciles» en función del promedio de aciertos obtenidos.

Por su parte, el análisis de los ítems que otorgaron mejoras postest para el grupo piloto de secundaria nos mostró, al igual que ya lo hiciera para el grupo de primaria, que la mayoría de los ítems de mejora habían sido catalogados como difíciles o muy difíciles. Concretamente, el 61% de los contrastes ganadores fue obtenido con este tipo de ítems. De esta forma, también este grupo habría tenido un comportamiento acorde con las hipótesis de la TRI, obteniendo sus mejoras en aquellos ítems que representaban una mayor dificultad inicial. Tal y como apuntábamos en el capítulo de Resultados, estos datos estarían avalando en cierto modo algunas de las más duras críticas que se han formulado en contra de la utilización de procedimientos convencionales de evaluación para analizar los estudios que desarrollan pruebas dinámicas. En este sentido, se vería reforzada la necesidad de utilizar procedimientos de análisis y evaluación complementaria, de corte cualitativo o mixto, que puedan aportar información adicional a la ofrecida por métodos de análisis exclusivamente psicométricos.

Por otra parte, los datos obtenidos del análisis de la evolución de las puntuaciones en la muestra de aplicación piloto, nos mostraron claramente ganancias significativas en el nivel de comprensión de textos, tanto en primaria como en secundaria. De esta forma,

quedaría verificada la sexta de nuestras hipótesis formuladas, relativa al progreso de aquellos sujetos sin especiales dificultades de lectura, en ausencia de intervención específica.

Con relación a los procesos de ajuste, en capítulos anteriores habíamos hecho ya referencia a las dificultades que entrañaba el cambio de actitudes. Apuntábamos que un cambio significativo en los procesos de ajuste personal-social, podría necesitar de intervenciones más específicas que la planteada a través del dispositivo de evaluación dinámica EDPL, y que quizás la duración de la aplicación experimental no resultara suficiente para provocar una transformación mensurable en este sentido. Hemos podido comprobar, sin embargo, que el grupo experimental obtuvo ganancias significativas en la escala de ajuste APSL. Estas ganancias se produjeron en función del segundo y del tercero de los criterios adoptados en nuestra investigación, es decir, en aquellos sujetos que fueron designados por el profesorado como alumnos con dificultades de aprendizaje de la lectura y que, además, obtuvieron bajas puntuaciones en la escala de ajuste en el pretest; y también en aquellos que, cumpliendo las dos condiciones anteriores, obtuvieron así mismo bajas puntuaciones en la prueba de CI. De esta forma, la aplicación del dispositivo EDPL habría provocado cambios significativos en los procesos de ajuste personal-social relacionados con la lectura, especialmente en aquel grupo de sujetos que más dificultades presentaba en este terreno.

El análisis de la evolución de las puntuaciones en APSL en la muestra piloto, podría reforzar los resultados mencionados anteriormente. En efecto, la no obtención de ganancias significativas por parte de este grupo de sujetos que no presentaba dificultades en su aprendizaje, vendría a confirmar que la modificación de actitudes requiere de intervenciones específicamente dirigidas a este fin. En este sentido, no se obtuvieron diferencias significativas en los contrastes efectuados para la muestra piloto —en contra de lo que sí sucedió en los niveles de comprensión—, ni para la puntuación global de la escala, ni para las distintas dimensiones. Por el contrario, la estabilidad observada en las puntuaciones pretest y postest, apunta en la dirección que mencionábamos más arriba.

Por otra parte, con relación a la posibilidad de que se hubiera producido un efecto *techo*, hemos comprobado en los análisis efectuados en páginas anteriores, que el subgrupo conformado por sujetos con privación socio-cultural, así como el integrado por aquellos que no precisaban apoyo a la integración ni presentaban discapacidad o privación, obtenían puntuaciones medias iniciales y finales similares a las obtenidas por el grupo piloto. Así mismo, para los grupos experimental y control, también algunas de las puntuaciones de las dimensiones contempladas en la escala, se asemejan a las obtenidas por la muestra piloto. Aunque ninguno de estos subgrupos mencionados incrementa significativamente sus puntuaciones en el postest, la existencia clara de un efecto *techo* hubiera ido acompañada de puntajes iniciales en el grupo piloto superiores en la fase inicial a los anteriormente mencionados, o de un incremento significativo de las puntuaciones en el grupo de tratamiento en la fase final del estudio, transcurridos seis meses. Por otra parte, hemos comprobado que la media de la puntuación directa del grupo piloto en la escala es de 55,49 puntos, todavía lejos de los 80 puntos posibles. Esto parece indicar la existencia de un amplio margen de mejora para este grupo, máxime cuando la mediana de la distribución es 55 y más del 35% de los sujetos obtiene más de 60 puntos, situándose la puntuación máxima alcanzada en 76 puntos.

A pesar de la igualdad reiterada entre las puntuaciones globales de los contrastes para las dos etapas, igualdad que se da también entre la mayoría de las dimensiones contempladas, es cuanto menos llamativa la diferencia de puntuación entre los sujetos de primaria y secundaria. En el análisis de la evolución de las puntuaciones obtenidas por el GE, podíamos observar que las medias pretest y postest para primaria se situaban en los 98 puntos, mientras que las medias para el grupo de secundaria descendían un punto con respecto a las anteriores, manteniendo en todo caso la igualdad entre pretest y postest. Esta consideración pone de manifiesto que el grupo piloto de primaria sí habría superado ampliamente al GE de esta misma etapa, obteniendo una diferencia de medias inicial y final de 9 puntos; sin embargo, el grupo piloto de secundaria ha obtenido una puntuación muy similar a la conseguida por el grupo experimental, con una diferencia entre sus medias de tan sólo 1 punto. Parece ser que la etapa educativa, asociada en este caso inextricablemente a la etapa psico-evolutiva de los sujetos, ha tenido un peso importante en el nivel de ajuste personal-social manifestado en la prueba. Ya comprobamos que el grupo de tratamiento conformado por los sujetos que partían de posiciones más bajas en los niveles de ajuste e inteligencia, obtenían ganancias significativas postratamiento, y que algunos subgrupos de aplicación en secundaria obtenían importantes diferencias entre sus puntuaciones iniciales y finales, lo que nos indica que también en esta etapa se han obtenido importantes ganancias tras la aplicación del tratamiento experimental. A pesar de esto, las puntuaciones han ido casi siempre por debajo de las obtenidas por los sujetos de primaria. Con relación a las causas de este fenómeno, apuntamos una posible explicación basada en los mayores niveles de adaptación escolar y de ajuste personal-social en el alumnado de primaria. Esto se concretaría, por ejemplo, en una percepción más positiva del alumnado más joven acerca del trato de los profesores o de la ayuda recibida, tanto del profesorado como de los familiares. Por otra parte, factores relativos al momento evolutivo de los sujetos de una y otra etapa, mediatizados por variables de índole contextual-educativo, podrían contribuir igualmente a la explicación de las diferencias observadas. Entre estos factores, pueden haber tenido una especial incidencia: (a) una percepción de uno mismo y de las propias capacidades e intereses en los más pequeños, que podría no ser del todo ajustada a la realidad; (b) una capacidad menor de reflexión y de conocimiento metacognitivo relacionado con la propia actitud, la motivación, los intereses o las expectativas; (c) un mayor grado de deseabilidad social en los más pequeños; (d) una mayor «sinceridad» en el alumnado mayor, unida a una mayor capacidad de análisis, reflexión y conocimiento sobre uno mismo, sus propias actitudes, motivación, intereses o expectativas.

10.5 *Discusión de los resultados en la evaluación complementaria*

Con relación al análisis sobre la evaluación complementaria efectuada, hemos destacado el uso conjunto de métodos cuantitativos y cualitativos para recabar información del proceso de aplicación del dispositivo. Los aplicadores han debido poner en práctica una serie de estrategias de análisis, que nos han permitido obtener información sobre la aplicación dinámica, al tiempo que realizar la valoración posterior. Entre éstas estrategias se encuentran: (a) la utilización de la observación durante el desarrollo de las sesiones de aplicación de EDPL; (b) el registro y posterior categorización de estas observaciones en las *Hojas de registro y observación*; (c) la valoración de la actividad de los sujetos, con base en los indicadores de evaluación previamente establecidos y recogidos en el Manual; (d) la discusión colectiva de los aplicadores en las sesiones del Grupo de Trabajo. Por otra parte, como expusimos al

exponer el análisis del proceso de aplicación, una vez finalizado éste, tuvo lugar la codificación y cuantificación de los valores otorgados a las observaciones efectuadas.

En línea con lo apuntado en el párrafo anterior, debemos destacar el análisis y la valoración cualitativa realizada de forma individual y colectiva en el Grupo de Trabajo, acerca del funcionamiento y la consecución de los objetivos del propio grupo; así como sobre el dispositivo dinámico. Este análisis nos ha aportado información muy valiosa procedente de los profesionales que han participado de forma directa en la aplicación de EDPL. En este sentido, podríamos destacar que el grado de satisfacción expresado por los integrantes del GT con relación al trabajo desarrollado ha sido muy alto, aunque en las valoraciones efectuadas quedaran igualmente recogidos algunos aspectos que incidieron negativamente en las posibilidades de una aplicación óptima del dispositivo. El proceso de valoración y análisis cualitativo efectuado, basado fundamentalmente en el análisis proposicional de las respuestas proporcionadas por los aplicadores a las cuestiones formuladas, nos ha permitido obtener información relevante conducente a la confirmación de los resultados mediante múltiples observaciones; así como a contrastar las distintas aproximaciones metodológicas llevadas a cabo para obtener los datos (Mora, 1998). Como ya apuntábamos en capítulos anteriores, y en este mismo, la consideración de diversos procedimientos de obtención de la información, así como de diferentes modelos de análisis de la misma, de carácter cuantitativo y cualitativo, habría configurado en nuestro estudio un proceso de *triangulación de la información*, que a su vez habría proporcionado un control complementario con relación a las diferentes fuentes de información empleadas (Lukas y Santiago, 2004).

De este modo, se ha tratado en todo momento de cotejar los datos provenientes de la evaluación complementaria con aquellos extraídos con base en la evaluación realizada mediante las pruebas de criterio. Uno de los aspectos fundamentales que debíamos evaluar consistía en la posible confluencia de resultados entre los datos provenientes de las distintas fuentes consideradas en el estudio, tanto aquellas más cuantitativas como las que nos ofrecieron sobre todo información cualitativa. Hemos analizado la consonancia entre los resultados obtenidos en las distintas pruebas criterio, junto con la valoración cuantificada del proceso de aplicación (puntuaciones dinámicas EDPL), la valoración del rendimiento escolar y del progreso de los sujetos, así como las opiniones de los aplicadores, expresada en el cuestionario de valoración del dispositivo. En este sentido hemos encontrado una mayor convergencia entre los datos provenientes de valoraciones con un carácter más procesual, es decir, de la valoración del proceso de aplicación dinámica, la valoración del rendimiento escolar y del progreso del alumnado, así como de las opiniones del profesorado aplicador. De forma más puntual, hemos encontrado confluencia de resultados entre algunos valores de las pruebas criterio y las puntuaciones dinámicas o la valoración del rendimiento escolar.

Además de los resultados, la convergencia se ha analizado con relación a otros aspectos centrales en nuestro estudio. Tal es el caso de la relevancia que tendrían los distintos procesos implicados en la lectura, en la explicación de las dificultades analizadas en nuestro estudio. Nuestra quinta hipótesis planteaba la relevancia de los procesos metacognitivos en la explicación de las dificultades relacionadas con la comprensión. La verificación de esta hipótesis se ha realizado a través del análisis de las puntuaciones obtenidas en las dimensiones metacognitivas de las pruebas ECO 1, ECOS y APSL, así como de las puntuaciones dinámicas obtenidas en EDPL, tanto las puntuaciones obtenidas de la valoración de los indicadores metacognitivos, como de la puntuación

total. Estos análisis debían mostrar, en primer lugar, que las puntuaciones metacognitivas obtenidas en las distintas pruebas analizadas, suponían una fuente importante de dificultades para el alumnado participante; en segundo lugar, debían mostrar que, si bien los procesos metacognitivos constituían efectivamente una fuente significativa de dificultades en la fase inicial del estudio, las puntuaciones referidas a su puesta en práctica deberían reflejar mejoras con el tratamiento en el grupo experimental. En este sentido, hemos encontrado convergencia entre los resultados obtenidos en las pruebas de comprensión y las puntuaciones dinámicas; en ambas pruebas hallamos unos niveles de puntuación que sitúan a los procesos metacognitivos entre los que representaron una mayor dificultad. Con relación a la muestra de primaria, los datos obtenidos de la prueba criterio, relativos a la localización de las mayores dificultades, apuntan en una dirección similar a la que podía derivarse del análisis del dispositivo de evaluación dinámica; sin embargo, más allá de la confirmación general de los resultados obtenidos, confirmación que se da sobre todo a nivel de *dónde* se encuentran algunas de las principales dificultades relacionadas con la lectura y la comprensión en sujetos con dificultades de aprendizaje, podría existir una escasa relación entre la información que puedan llegar a aportar pruebas estáticas y dinámicas con relación a determinados procesos implicados en la lectura; estas limitaciones observadas estarían relacionadas con: (a) la capacidad diferencial de ambos modelos para profundizar y afinar en el análisis de *dónde* se encuentran las dificultades de lectura y de por qué se producen; (b) la capacidad diferencial para explicar el proceso mismo aprendizaje (integración de modelos de acción, incorporación de estrategias, etc.); (c) la capacidad diferencial para facilitar y explicar un mayor rendimiento en las actividades desarrolladas mediante la aplicación dinámica.

En cuanto a secundaria, el análisis de las mejoras con relación a la convergencia inicial observada, ofrece resultados positivos para el grupo experimental. Este grupo experimenta ganancias significativas en sus puntuaciones metacognitivas de la prueba de comprensión, desapareciendo estas dimensiones, en la fase posttest del estudio, de entre las que muestran un mayor grado de dificultad. En efecto, pudimos comprobar que las puntuaciones promedio que indicaban el grado de dificultad que habían representado los distintos procesos evaluados en la prueba de comprensión utilizada como criterio, se veían sensiblemente incrementados de forma significativa en la fase posttest. De este modo, la dimensión metacognitiva *Supervisión y autoevaluación del proceso de comprensión*, desaparecía de entre las más difíciles a tenor de las puntuaciones obtenidas tras el tratamiento. Apuntábamos cuando abordábamos la exposición de los resultados, que el trabajo específicamente dirigido hacia las habilidades de control y regulación consciente de los propios procesos de comprensión, a lo largo de las sesiones de aplicación del dispositivo, habría estado en la base de la evolución positiva de las puntuaciones en las dimensiones metacognitivas de la prueba. Estas ganancias significativas no se produjeron en el grupo de control, para el que, obviamente, tampoco se produjo evolución en la consideración del grado de dificultad que representó en el posttest la dimensión metacognitiva de la prueba.

Los datos obtenidos también nos ofrecen otra información que, sin duda, es relevante en función de los objetivos de nuestro estudio. En este sentido, pudimos comprobar que las puntuaciones dinámicas obtenidas por el grupo experimental, superaban de forma significativa a las puntuaciones obtenidas en las pruebas de comprensión, conseguidas en la fase pretest. Esto ocurría tanto para el alumnado de primaria como para el de secundaria, y en todas las dimensiones contempladas en el dispositivo, además de en la

puntuación dinámica global. Apuntábamos que, a pesar de que ambas puntuaciones no pueden compararse directamente, sí resulta de nuestro interés constatar que los resultados obtenidos por el alumnado del grupo de tratamiento en una situación de aprendizaje en colaboración, son mejores que los resultados en una prueba de carácter estático. De este modo, la valoración cuantificada del proceso de aprendizaje ha tenido valores más altos que la valoración del producto de la comprensión, efectuada a raíz de la aplicación de las pruebas criterio de comprensión. Las situaciones de aprendizaje mediado o en colaboración, que han tenido lugar durante las sesiones de trabajo con el dispositivo EDPL, habrían propiciado de este modo, mayores posibilidades de evaluación en la *Zona de Desarrollo Próximo*, acerca de aquello que el sujeto es capaz de realizar con la ayuda ajustada, así como de detectar con mayor precisión dónde se encuentran sus dificultades y cómo podríamos optimizar su rendimiento. Por tanto, los datos obtenidos nos hablarían de actividades realizadas por los sujetos en situaciones de aprendizaje mediado, que fueron valoradas positivamente, con lo que, en función de la teoría socio-cultural del aprendizaje (Vygotski, 1934/1995), cabría esperar que estos mismos sujetos pudieran realizar posteriormente estas mismas actividades de forma autónoma. Ya mencionamos que, en parte, este hecho ha podido ser verificado con base en las mejoras significativas obtenidas por el grupo experimental en función de los criterios adoptados.

Por otra parte, el análisis gráfico de los resultados obtenidos en las distintas dimensiones del dispositivo, nos mostraba una cierta tendencia descendente en las puntuaciones promedio, obtenidas a medida que se aplicaba EDPL. Esta tendencia era más acusada en algunas de estas dimensiones, entre ellas la que hacía referencia a los procesos metacognitivos. Hemos tratado de encontrar explicación con base en varios motivos. En primer lugar, podemos comprobar que, en general, las valoraciones de los indicadores pertenecientes a un determinado proceso, han sido más altas en las actividades específicas pertenecientes a ese mismo bloque; probablemente, el aplicador ha prestado una mayor atención a la evaluación de los indicadores de los procesos más relacionados con el bloque de actividades que estaba realizando en ese momento. En la misma dirección apuntaría el hecho de haber dejado sin valorar algunos indicadores, debido a que no se realizaron algunas de las actividades propuestas; esto habría ocasionado en general una cuantificación mínima (1 punto), bajando sensiblemente el promedio. En este sentido, el bloque de actividades de procesos metacognitivos no pudo ser finalmente realizado, tal y como referimos anteriormente; es posible por tanto que, siguiendo la tendencia apuntada de mayores puntuaciones en los indicadores pertenecientes al bloque de procesos específico que se realizaba, se hubiera producido un repunte de las puntuaciones metacognitivas de haber sido realizadas las tres actividades específicas propuestas en EDPL.

Otro aspecto de interés, relacionado en parte con las tendencias observadas, es que objetivamente la dificultad de las actividades ha ido en aumento a medida que avanzaba la aplicación. En efecto, el procedimiento de aplicación propuesto en el diseño de investigación contemplaba una realización secuenciada, que conllevaba una progresiva dificultad y complejidad de las tareas a ejecutar. De esta forma, al comienzo estaban situadas las actividades específicas de los procesos psicológicos subyacentes; posteriormente las pertenecientes al bloque de procesos de asociación G-F, seguidas por las actividades de los procesos de integración textual, integración texto-conocimientos previos, procesos metacognitivos y, por último, procesos de ajuste P-S. Los aplicadores coincidieron ampliamente en sus valoraciones acerca de la dificultad progresiva que

representaban las actividades para el alumnado participante. En efecto, las primeras actividades contenían básicamente tareas lingüísticas y no lingüísticas relacionadas con habilidades de orientación espacial, secuenciación temporal, memoria, lectura de pseudopalabras, conciencia fonológica y decodificación; representando, como hemos tenido oportunidad de comprobar, una dificultad seria para un pequeño número de sujetos. De esta forma, muchos de los indicadores propuestos fueron observados por los aplicadores y valorados positivamente. Sin embargo, a medida que avanzaba la aplicación del dispositivo, se iban introduciendo tareas que exigían la puesta en práctica de habilidades de lectura y comprensión de textos de progresiva dificultad, con requerimientos de integración del contenido del texto con los conocimientos y experiencias previas, realización de inferencias, distinción de la información relevante, etc. En este contexto, un número significativamente mayor de alumnos y alumnas encontró dificultades. Este hecho, lógicamente, habría tenido influencia en la menor valoración de los indicadores propuestos, realizada por parte de los aplicadores, e incluso en su no aparición en el transcurso de la tarea, con lo que su valoración cuantitativa habría sido mínima. Podemos observar, por ejemplo, en la gráfica de los procesos de integración textual, que la línea que marca la evolución de las puntuaciones desciende bruscamente hasta su nivel mínimo; en este caso, todos los sujetos que realizaron la actividad 28, perteneciente al bloque de procesos de integración texto-conocimientos, y considerada especialmente difícil por los aplicadores que la realizaron, obtuvieron una puntuación mínima, debido a que los indicadores de los procesos de integración textual no fueron observados ni valorados por los aplicadores.

Discusión del análisis del dispositivo en función del perfil de cuatro puntos

En nuestra valoración del dispositivo EDPL y de los resultados obtenidos con su aplicación experimental, hemos adoptado un sistema de análisis propuesto por Sternberg y Grigorenko (2003). Este análisis considera *cuatro puntos* o elementos clave que deben ser evaluados en la implementación experimental de una prueba de evaluación dinámica. En este sentido, tal y como expusimos durante el desarrollo del capítulo de Evaluación complementaria, hemos adaptado a nuestras necesidades y circunstancias, los elementos clave de la propuesta. De este modo, se han analizado fundamentalmente los tres primeros puntos, derivando el cuarto a otros investigadores que consideren relevante efectuar replicaciones futuras o revisiones de nuestro estudio, con el fin de evaluar la solidez de los resultados obtenidos. El primer aspecto analizado hacía referencia a la *información comparativa* del dispositivo dinámico. Hemos efectuado una doble valoración diferencial del modelo propuesto; en primer lugar, con relación a otros modelos dinámicos de los descritos y analizados en nuestro propio trabajo; en segundo lugar, con relación a pruebas convencionales que evaluaban el mismo constructo que EDPL.

Con relación a la primera de estas valoraciones comparativas, es relevante volver a destacar aquí que la aplicación de EDPL no ha supuesto un entrenamiento en los elementos pertenecientes a una determinada prueba que sería evaluada posteriormente. Nuestra propuesta de evaluación dinámica incluye esencialmente actividades que están estrechamente relacionadas con contenidos presentes en el contexto educativo y social. Concretamente, se proponen tareas de evaluación de la lectura y la comprensión de textos, con contenidos que se corresponden mayoritariamente al área de Lengua. Esta opción ha contribuido significativamente a que los resultados obtenidos adquieran un mayor grado de validez ecológica. Los contenidos trabajados durante la aplicación

dinámica pueden ser relacionados con los contenidos escolares, con lo que las predicciones efectuadas sobre el rendimiento académico podrían estar sólidamente fundadas. Así mismo, hemos podido valorar el proceso de aprendizaje de cada sujeto participante en actividades contextualizadas, que representaban grados de dificultad variables, y hemos deducido del proceso de aplicación una serie de patrones metodológicos que se mostraron eficaces en la mejora de las habilidades evaluadas, o bien en la obtención de información valiosa. Hemos tratado, por tanto, de otorgar una especial relevancia a la contextualización de los instrumentos de evaluación, en contraposición a algunas de las propuestas dinámicas que tuvimos la oportunidad de revisar y analizar (Tzuriel, 1999, 2000; Fernández-Ballesteros y Calero, 2000). Esta consideración de la relevancia de la contextualización del proceso de evaluación y de los instrumentos, viene determinada con relación al establecimiento de mayores índices de validez predictiva del dispositivo; y, sobre todo, a efectos de poder ofrecer orientaciones relativas a la optimización del proceso de enseñanza y aprendizaje. En este sentido, ya hicimos referencia, cuando analizábamos los estudios y la propuesta de de ED de Fernández-Ballesteros y sus colaboradores, al uso preferente que se había hecho de pruebas descontextualizadas con el objetivo primordial de minimizar la influencia de la variable «historia previa de los sujetos». En nuestra revisión crítica del modelo, habíamos expuesto, en cambio, que la descontextualización del contenido de las pruebas, podía llegar a tener una relevancia notable con relación a las posibilidades de generalización de los conocimientos y las estrategias trabajados. Se ha descrito en numerosas ocasiones que, en el contexto escolar, los contenidos de carácter específico juegan un papel relevante en la enseñanza y aplicación de estrategias para la mejora del aprendizaje (Gardner, 1983; Chi, 1987; Carey y Spelke, 1994; Martín, 1999; Jensen, 2000; Mateos, 2001; Navarro, 2004).

Así mismo, poníamos en duda la efectividad y la conveniencia de tratar de eludir la historia educativa del sujeto y el contexto sociocultural en el que tienen lugar los aprendizajes que se quieren evaluar. Bajo nuestro punto de vista, la utilización de instrumentos de evaluación con contenidos descontextualizados, no sería imprescindible en un proceso dinámico de evaluación de la capacidad de aprendizaje. Éste proceso podría tener lugar mediante el uso de actividades contextualizadas y dirigidas de forma especial a la evaluación de los procesos metacognitivos implicados, realizadas en situaciones de colaboración y de interacción mediada, situados en la *zona de desarrollo potencial* del sujeto, con la intención de determinar dónde se encuentran sus dificultades y cómo podríamos optimizar el proceso de aprendizaje. Apuntábamos igualmente las dificultades que conllevaría la asunción de un paradigma más cualitativo para establecer un control adecuado de las variables, al tiempo que índices óptimos de validez y fiabilidad. Somos conscientes, por otra parte, de los problemas que supone la concreción de una propuesta de ED que contemple la contextualización del proceso y de los instrumentos de evaluación, y que al mismo tiempo considere las medidas cuantitativas como los principales criterios de valoración de los resultados. En este sentido, proponíamos una serie de claves que podían minimizar estas dificultades, permitiendo afrontar, de este modo, una evaluación dinámica más contextualizada. A pesar de que el formato de nuestra propuesta dinámica no sea el de pretest /mediación/ postest, la revisión de estas claves nos sitúa en la aplicación experimental de nuestro dispositivo EDPL. En efecto, al contar con un pretest y un postest, con el propósito ya analizado de obtener un grado de control mayor de las diversas variables que podían distorsionar los resultados en una investigación como la nuestra, hemos puesto en práctica, en mayor o menor medida, las distintas claves apuntadas. De esta forma, se

han determinado de forma exhaustiva y previa, mediante la pasación de pruebas específicas, los conocimientos de los sujetos sobre el contenido concreto objeto del estudio, es decir, la comprensión de textos; se ha desarrollado un proceso de mediación, encaminado especialmente a valorar y facilitar la puesta en práctica y el aprendizaje de aquellos procesos cognitivos y metacognitivos que están implicados de forma significativa en el desarrollo y dominio de las actividades planteadas; se ha determinado una medida del cambio producido con respecto a la línea base establecida, y se ha obtenido esa información con la pasación de una prueba postratamiento equivalente a la inicial; y, por último, se ha contado, en general, con grupos de control adecuados, que nos han permitido validar unas conclusiones positivas en cuanto a la mejora en los procesos de comprensión de los sujetos, así como en los procesos de ajuste y en los niveles de rendimiento cognitivo, al tiempo que nos han permitido extraer información muy valiosa con relación a la mejora del proceso mismo de evaluación.

Nuestra propuesta de ED ha contemplado el análisis del proceso de aprendizaje de un sujeto en una situación de interacción mediada, mediante el desarrollo de actividades que se sitúan en su ZDP con el fin de determinar el potencial dinámico de aprendizaje, estableciendo aquellas medidas de apoyo necesarias para optimizar el proceso de enseñanza y aprendizaje, favoreciendo experiencias de aprendizaje mediado que sean significativas para el sujeto y, así mismo, el acceso progresivo de éste a formas de funcionamiento cognitivo más autónomas y eficaces. La asunción de un paradigma de corte más cualitativo y orientado a la mejora del proceso de enseñanza y aprendizaje, parece más acorde con estas prerrogativas, aunque es importante destacar que también para nosotros constituye un objetivo fundamental de la investigación dinámica la continua mejora de los aspectos relacionados con la validez y la fiabilidad de las pruebas y los procedimientos.

Por otra parte, con relación a la segunda valoración diferencial efectuada, hemos utilizado diversos indicadores: la calidad de los datos obtenidos y la fiabilidad de las valoraciones efectuadas en la aplicación de EDPL; el tipo de información ofrecida; y las aportaciones metodológicas suplementarias a la aplicación de pruebas de carácter estático. En este sentido, debemos hacer referencia inicialmente a ciertos problemas que se observaban en algunas de las propuestas de ED analizadas; algunos de los cuales pueden ser compartidos, en mayor o menor medida, por nosotros. Comentábamos por ejemplo, que en el LPAD de Feuerstein, la fiabilidad de la fase de mediación podía constituir un serio problema, por cuanto la relevancia del papel del evaluador en el proceso, podría dificultar la replicabilidad de las observaciones y valoraciones realizadas, exigiendo una formación exhaustiva de los profesionales que aplicaran el dispositivo (Lidz, 1991; Tzurriel y Samuels, 2000; Lidz y Elliott, 2000; Sternberg y Grigorenko, 2003). Así mismo, otras observaciones críticas al modelo de Feuerstein, aludían a la «vaga» definición de los conceptos relacionados con la teoría de base propuesta. Los problemas para establecer conexiones entre el marco teórico propuesto y los hallazgos de la evaluación, dificultarían el establecimiento de validez en los resultados (Bücheel y Scharnhorst, 1993); así como la posibilidad de operativizar cuantitativa o cualitativamente los resultados obtenidos en el proceso de evaluación. Con relación al primero de los problemas planteados, hemos hecho mención en el transcurso de este trabajo, a la relevancia del proceso de formación del profesorado en la aplicación de EDPL, así como a las dificultades que, en la propia aplicación experimental del dispositivo, ha supuesto conjugar una implementación contextualizada y flexible de la propuesta, con la necesidad de establecer parámetros sólidos de validez

y fiabilidad. Posteriormente, profundizaremos en algunos aspectos relacionados con la fiabilidad de la fase de mediación, cuando analicemos algunas de las valoraciones y propuestas de mejora efectuadas por los aplicadores en las sesiones de evaluación conjunta del dispositivo. Así mismo, con relación al segundo de los problemas planteados, hemos hecho referencia; tanto en el análisis de los resultados obtenidos en la batería criterio, como en la evaluación complementaria, a las conexiones relevantes y significativas entre los resultados obtenidos y los distintos procesos implicados en la lectura que han sido contemplados en nuestra teoría de base en esta investigación.

Por nuestra parte, ya hicimos mención, con relación a la calidad de los datos obtenidos, a que habíamos optado por realizar una estimación de los índices de fiabilidad en función de la valoración diferencial efectuada por los distintos aplicadores. Tratábamos de este modo, de responder a las exigencias expresadas por Sternberg y Grigorenko (2003), para quienes el reto fundamental para las pruebas dinámicas lo constituía la fiabilidad que podían mostrar frente a la valoración realizada por distintos evaluadores sobre los elementos clave contemplados en la propuesta. En nuestro caso, las valoraciones diferenciales se han producido de diversas formas y en distintos momentos. Por una parte, mediante el análisis cualitativo de los aplicadores, efectuado a través de la discusión colectiva en las sesiones del Grupo de Trabajo, así como mediante las respuestas otorgadas a las preguntas planteadas en el cuestionario sobre la aplicación experimental del dispositivo EDPL; por otra parte, a través de su valoración cualitativa del desempeño de los sujetos en el transcurso de las actividades; valoración que quedó plasmada en las Hojas de registro y evaluación, y que fue posteriormente cuantificada en función de los criterios ya comentados y analizados con anterioridad. Es esta última valoración la que nos ha permitido aproximarnos al grado de fiabilidad del conjunto de las observaciones y valoraciones efectuadas. En este sentido, la estimación de la fiabilidad inter-observadores, mediante el cálculo del coeficiente de correlación intra-clase, nos permitió constatar que el nivel de consistencia interna de las observaciones efectuadas ha sido especialmente alto, obteniéndose valores cercanos al acuerdo absoluto en la estimación realizada con el total de indicadores de evaluación valorados en las actividades, así como también en la estimación efectuada con los indicadores de procesos metacognitivos.

Un segundo aspecto analizado ha sido el *valor de predicción* del dispositivo EDPL, incidiendo especialmente en el valor predictivo con relación al rendimiento académico, así como con relación a las pruebas criterio de comprensión, ajuste personal-social e inteligencia. Por último, hemos analizado el *grado de eficacia* del dispositivo en función del tiempo y el esfuerzo necesarios para obtener la información deseada. Dentro de este último apartado valorado, hemos recogido un análisis global sobre la eficacia de EDPL en función de los diversos indicadores ya mencionados en los puntos anteriores. Cabe preguntarse con relación al tiempo y al esfuerzo que es necesario invertir para poder obtener información considerada útil y eficaz en función de los objetivos planteados, si el dispositivo de evaluación dinámica que proponemos en este trabajo, puede ser llevado al contexto educativo para su utilización en la práctica por parte de los profesionales para los que ha sido pensado. En este sentido, estamos de acuerdo con Saldaña (2001) en que se hace necesario, por una parte, mejorar continuamente el proceso de evaluación propuesto, con el fin de optimizar los esfuerzos y el tiempo que es necesario emplear en su aplicación en contextos de interacción; por otra parte, consideramos también ineludible una llamada a la reflexión dentro del ámbito escolar sobre los objetivos de la evaluación de la comprensión de textos, y sobre si es posible

acceder a la información deseada y útil para la mejora del proceso, con prácticas que conllevan escaso tiempo y nivel de formación, al tiempo que se caracterizan por una cierta simplicidad de los análisis efectuados. Volvemos a destacar aquí la preeminencia de la obtención de información cualitativamente idónea con vistas a la mejora del proceso de enseñanza y aprendizaje, sobre el tiempo y la energía empleados para este fin. Como ya pusimos de manifiesto a lo largo de nuestro trabajo, la flexibilidad de nuestra propuesta original en cuanto al tipo y número de actividades que pueden ser puestas en práctica durante la aplicación del dispositivo, comporta la relatividad de la dimensión temporal en función de las necesidades de información sobre los procesos implicados en las dificultades de lectura, o sobre los patrones metodológicos de mediación que se consideran óptimos para la mejora. De este modo, tiempo de aplicación y formación necesaria de los aplicadores, se tornan conceptos inseparables si se quiere optimizar el proceso de aplicación, obteniendo el máximo de información cualitativamente relevante, y adicional a la que sería posible obtener con la aplicación de pruebas más convencionales.

Con relación al valor de predicción del dispositivo sobre el rendimiento en comprensión para los sujetos de primaria, los resultados parecen confirmar, en línea con lo que hemos argumentado con anterioridad, que las puntuaciones dinámicas, obtenidas de la valoración de un proceso de aprendizaje, se relacionarían mal con puntuaciones obtenidas a través de pruebas de corte estático, y centradas sobre todo en el producto final de la comprensión. En este sentido, con relación a esta ausencia de relación entre las puntuaciones dinámicas y las puntuaciones obtenidas en las pruebas de comprensión, podría tener lugar el hecho, destacado por Saldaña (2001), de que los procesos de evaluación dinámica aportarían información más valiosa sobre la construcción y el proceso de unas determinadas funciones (metacognitivas o de otra índole), que sobre una medida de estas funciones ya «cristalizada». Por otra parte, el hecho de que la correlación entre las puntuaciones EDPL y el pretest resulte negativa, siendo la puntuación dinámica significativamente mayor que la puntuación pretest, puede entenderse como un resultado a favor de la primera hipótesis experimental, ya que revelaría mejoras en los procesos evaluados, que no se recogen en los resultados de la fase inicial. Igualmente, esta diferencia podría matizar la ausencia de ganancias significativas en el GE entre las puntuaciones pretest y postest de la prueba ECO 1. Por otra parte, podemos observar que, a pesar de que la correlación de la puntuación EDPL con el postest es también negativa, el valor es sensiblemente menor que la obtenida con el pretest, por lo que podría entenderse que la divergencia entre las valoraciones habría descendido. Estos resultados se encuentran en sintonía con los obtenidos por Guthke y Wingenfeld (1992), quienes también usaron la valoración del profesorado y las calificaciones escolares como medida del rendimiento escolar. Estos autores comprobaron que las puntuaciones postest de las Pruebas de Aprendizaje correlacionaban significativamente más alto con el rendimiento escolar que las obtenidas en la fase pretest. En cualquier caso, la correlación negativa de la puntuación dinámica con la puntuación postest, dado que no existían diferencias significativas entre ésta y la obtenida en el pretest, observándose sin embargo estas diferencias con relación a los puntajes de EDPL, apunta nuevamente a que las valoraciones efectuadas sobre el proceso de aprendizaje han ido más allá de las observadas en las pruebas criterio, indicando ganancias en los procesos trabajados que no se reflejan en las pruebas estáticas. Podríamos decir, tal y como se desprende de los estudios de Beckmann y Guthke (1999), que la información obtenida de la valoración del dispositivo dinámico, incluiría la información que se podría obtener con la pasación de las pruebas estáticas de

comprensión o de ajuste; y que, por el contrario, los datos que éstas pueden proporcionar, no nos ofrecerían la información que se podría obtener de la aplicación de EDPL. Por otra parte, también Swanson (2000) encuentra entre los datos de sus estudios sobre el potencial de procesamiento cognitivo, que algunos de los componentes del TPC-S de corte metacognitivo, no se hallaban positivamente correlacionados con medidas convencionales y estáticas del mismo constructo. En efecto, las medidas dinámicas relacionadas con procesos de conocimiento metacognitivo o autorregulación, encontrarían débiles correlaciones con medidas estándar. Por otra parte, la falta de sintonía entre nuestras puntuaciones dinámicas y las obtenidas en las fases pretest y postest, se torna en una mayor afinidad del valor de correlación de la puntuación EDPL con la mejora; puntuación ésta última que remite a un proceso que tiene lugar entre dos aplicaciones estáticas. En efecto, el valor de correlación pasa de ser negativo con las dos puntuaciones estáticas, a tener un valor positivo de $r=0,228$. Este incremento significativo en los valores de correlación entre la mejora y la puntuación obtenida de la aplicación de EDPL, se produjo tanto en la puntuación dinámica global como en la puntuación dinámica referente a los procesos metacognitivos.

Otros datos reveladores son aquellos que tienen que ver con las calificaciones finales otorgadas al alumnado participante, así como con los valores de correlación que se establecieron entre las distintas puntuaciones contempladas y las calificaciones. En efecto, el uso de las calificaciones escolares como criterio de rendimiento académico, merece una mención aparte. Ya hicimos alusión a que el valor de predicción del dispositivo EDPL se establecería mediante dos criterios externos: por una parte, la valoración del profesorado acerca del rendimiento escolar en el área de Lengua, mediante la cumplimentación de una plantilla en la que, para cada alumno, el docente debía asignar un determinado valor en función del nivel manifestado por éste en cada uno de los siete criterios de evaluación contemplados; y, por otra, las calificaciones escolares en el área de Lengua al finalizar el curso. A pesar de que conocíamos que el uso de este segundo criterio planteaba inicialmente algunos problemas de fondo, decidimos recabar toda la información posible con el fin de enriquecer y matizar los resultados que pudiéramos obtener de las valoraciones. En efecto, de todos es sabido que los criterios de evaluación utilizados en los Centros para establecer la calificación final del alumno, son muy dispares. Esta disparidad de criterios se agudiza cuando se trata de alumnado que presenta necesidades educativas especiales o dificultades de aprendizaje. En algunos Centros se opta por un criterio que podríamos denominar homogeneizador, fundamentado en la valoración académica del alumnado con necesidades educativas especiales con relación a los criterios que rigen para todo el alumnado. Sin perjuicio de que la información que se traslada a las familias pueda contener las oportunas observaciones sobre el progreso de los alumnos, o de que se adjunten los informes del profesorado de apoyo a la integración sobre la marcha de los mismos; en ausencia de una ACI (Adaptación Curricular Individualizada) que marque los criterios de evaluación adaptados, la valoración final que se hace del alumno en los boletines de notas, refleja la no consecución de los objetivos mínimos del curso donde se encuentre, aunque pueda ser evidente que se ha producido un progreso significativo en el aprendizaje del alumno. Por tanto, las calificaciones obtenidas de los alumnos que pertenecen a estos Centros, poco tienen que ver con el progreso real que hayan podido obtener en su comprensión. Por otra parte, en otros Centros, se actúa de forma casi opuesta, otorgando una valoración positiva por defecto al alumnado con NEE, entendiendo que *progresan según su capacidad o según sus posibilidades*. En estos casos, puede ser que la valoración académica que se haya reflejado finalmente en los

boletines, tampoco se haya visto relacionada con el progreso del alumno en los criterios del área de Lengua.

Lógicamente, sólo en aquellos casos en los que el criterio adoptado por el Centro haya estado en función de una adaptación real de los criterios de evaluación y, por tanto, el alumno haya sido valorado finalmente en función del progreso obtenido con relación a los objetivos marcados para él, la correlación de las calificaciones en el área de Lengua con los demás valores observados, ha debido ser positiva. Sin embargo, tanto el análisis de datos efectuado como la información otorgada por los profesores aplicadores, nos mostraron que efectivamente existía una notable divergencia entre los criterios de evaluación adoptados por los distintos Centros colaboradores. Los datos analizados muestran que estas divergencias hacen muy dudosa la validez del criterio de rendimiento académico basado en las calificaciones escolares, al menos en educación primaria.

Pudimos comprobar que la única correlación positiva que establecía las calificaciones finales, tenía lugar con las puntuaciones pretest. En este sentido, una posible explicación a estos datos proviene de la posible falta de relación existente entre las calificaciones y el rendimiento real de los sujetos, debido a las razones que hemos ya argumentado; sin embargo, otro posible análisis, que puede ser complementario al primero, apuntaría a que las calificaciones finales habrían reflejado el «estado inicial» del sujeto, sin que los cambios producidos se hayan visto reflejados en las valoraciones académicas otorgadas al finalizar el curso. En síntesis, el análisis de los datos relativos al rendimiento escolar en primaria, basado en las calificaciones escolares, otorga poca consistencia al criterio basado en estas valoraciones. La divergencia entre la propia valoración del tutor sobre el rendimiento del alumno y las calificaciones académicas finales, muestra claramente la falta de consistencia de este criterio. Si a esto unimos los datos sobre la relación entre las calificaciones y los demás valores observados, así como la información acerca de la diversidad de criterios adoptados en los Centros para la valoración del alumnado con NEE, hemos de concluir que el criterio de rendimiento escolar basado en las notas no parece cumplir en esta etapa la función para la que ha sido incluido en el diseño de investigación, y que, por lo tanto, debe ser tomado en cuenta con muchísima cautela.

Otro aspecto a valorar, de suma importancia, es la puntuación de mejora. Tal y como pudimos comprobar en el capítulo dedicado a la evaluación del proceso de aplicación del dispositivo, esta puntuación mantenía un valor de correlación positiva con la puntuación dinámica, al igual que con la puntuación posttest, llegando a ser en este caso muy significativa; por su parte, también mantenía correlación positiva con la valoración del tutor sobre el rendimiento escolar, así como con la valoración del profesorado de apoyo sobre el rendimiento y sobre el progreso de los sujetos, ambas también significativas. En cambio, la correlación fue negativa y significativa con el pretest, con la valoración del progreso expresada por los tutores y con las calificaciones escolares. Del análisis de estos datos, podemos desprender algunos elementos relevantes para nuestra discusión. Por una parte, la valoración del progreso efectuada por el profesorado de apoyo está más en consonancia con la puntuación que refleja las diferencias observadas entre pretest y posttest, que la valoración del progreso efectuada por el profesorado tutor. Por otra parte, es significativa la divergencia entre los contrastes de la puntuación de mejora con las puntuaciones inicial y final; en este sentido, y teniendo en cuenta que no se hallaron diferencias estadísticamente significativas entre ambas puntuaciones para este grupo, los datos apuntan a que las puntuaciones obtenidas en el

pretest no predicen en absoluto la dirección de las diferencias observadas entre el pretest y el postest. Dado que el pretest correlaciona también negativamente con la puntuación dinámica, podemos concluir que la puntuación obtenida inicialmente en una prueba estática no predice la actuación en el dispositivo de evaluación dinámica, ni asimismo los cambios observados tras su aplicación con respecto a los resultados de la misma prueba realizada posteriormente. En este sentido, estos datos vendrían a demostrar que, con relación a las dos puntuaciones establecidas como indicadores de la actuación en el dispositivo EDPL, es decir, la puntuación dinámica y la puntuación de mejora, el pretest otorgaría una información a todas luces insuficiente sobre los sujetos. En palabras de Cioffy y Carney (1983), las pruebas de lectura que evalúan el producto de la comprensión, y que son administradas convencionalmente, podrían aportar información sobre los conocimientos del alumno y sobre las habilidades de comprensión; pero no pueden llegar a otorgar información sobre el potencial de aprendizaje, y, sobre todo, suelen ser de poca ayuda a la hora de identificar las condiciones optimizadoras del proceso. De este modo, los datos obtenidos estarían en consonancia con la mayoría de los que se aportaban en nuestra revisión teórica de los distintos modelos de ED (Por ejemplo: Budoff, 1987a; Campione y Brown, 1987; Guthke y Beckman, 2000; Resing, 2000; Sternberg y Grigorenko; 2003).

Con relación a los valores obtenidos por el GC, hay que resaltar las menores puntuaciones obtenidas en todas las valoraciones efectuadas por los tutores y por el profesorado de apoyo, sobre el rendimiento escolar y sobre el progreso del alumnado. Estas valoraciones estuvieron sensiblemente por debajo de las que recibieron los sujetos que participaron del grupo experimental. Por otra parte, los resultados obtenidos por el GC con relación a los valores de correlación entre las calificaciones escolares y las puntuaciones postest, así como entre las calificaciones y la valoración del tutor o las puntuaciones de mejora, parecen apuntar en la dirección que comentábamos cuando analizamos la composición de los grupos completos experimental y control. En este sentido, la mayor presencia en el GC de alumnado sin discapacidad y que no recibía apoyo a la integración, podría explicar, al menos en parte, que sus calificaciones escolares hayan estado más acordes con su rendimiento real en el área de Lengua, sin que se haya producido mayoritariamente —como todo hace indicar que ha ocurrido en el grupo experimental— la divergencia a la que antes aludíamos entre el rendimiento y los criterios de evaluación adoptados para valorar este rendimiento. De esta forma, en el GC se da el resultado natural de correlación positiva y significativa entre la valoración del tutor sobre el rendimiento en Lengua y las calificaciones finales en este mismo área.

Los resultados de los análisis de regresión efectuados, nos mostraron que la puntuación dinámica predecía de forma significativa la mayoría de los valores analizados. En este sentido, su inclusión en la ecuación del modelo como única variable explicativa, tenía lugar para el rendimiento escolar valorado por los tutores y para la valoración del progreso del alumnado, valorado por el profesorado de apoyo. También se producía su inclusión en el modelo para la predicción del progreso de los sujetos, valorado por los tutores, así como para la predicción de las puntuaciones de mejora. En ambos casos, la puntuación dinámica contribuía de forma significativa a la explicación de la varianza observada. En el caso de la puntuación de mejora, ya apuntamos que la determinación de las puntuaciones que mejor explicaran este valor, podría aportarnos información sobre la validez diferencial de las puntuaciones dinámicas sobre las estáticas. En este sentido, pudimos comprobar que mientras el pretest correlacionaba negativamente con la mejora, la puntuación dinámica era incluida en la ecuación de regresión junto con el

postest, incrementando de forma significativa el valor de ajuste del modelo. En cuanto al rendimiento académico valorado por el profesorado de apoyo, era la puntuación de mejora la que mejor explicaba este valor, aunque la puntuación dinámica volvía a actuar como variable predictora, explicando por sí sola y de forma significativa la varianza observada. Por último, sólo la puntuación postest, que fue predicha mejor por la puntuación de mejora que por el pretest, y las calificaciones escolares, para las que ninguna variable fue incluida en la ecuación de regresión, revelaron la ausencia de significatividad de los valores de correlación obtenidos por las puntuaciones dinámicas.

Los análisis de regresión efectuados para los grupos de mejoradores y no mejoradores, revelaron, en ambos grupos, que eran también las puntuaciones dinámicas las que lograban explicar mejor la valoración del rendimiento académico efectuado por los tutores. De esta forma, la puntuación dinámica aparecía como el valor que mejor explicaba y predecía el rendimiento escolar. Los datos nos informaron también sobre la menor valoración promedio recibida por el grupo de no mejoradores, lo que sitúa la valoración efectuada por los tutores, en consonancia con las valoraciones realizadas por los aplicadores sobre el proceso de aprendizaje durante la aplicación del dispositivo.

Con relación al profesorado de apoyo, pudimos constatar que, al igual que ocurriera con la valoración de los tutores, la valoración media del rendimiento escolar de los no mejoradores descendió notablemente con respecto a la de los mejoradores. Estos resultados contribuyen a la validación del procedimiento utilizado para evaluar el rendimiento académico, ya que, tanto el profesorado tutor como el de apoyo, han realizado sus valoraciones acorde con los resultados que se han observado en los sujetos experimentales; en este caso, la diferencia también es muy significativa entre mejoradores y no mejoradores (grupo «ampliado») para la puntuación dinámica. Recordemos que la valoración del profesorado de apoyo coincide con la de los profesores-aplicadores cuando hablamos del GE, por lo que se establece una relación directa entre la catalogación de no mejoradores, en función de su actuación en las pruebas pretest y postest, y la valoración que los aplicadores efectúan sobre su actuación en el dispositivo EDPL, sensiblemente por debajo de la que reciben sus compañeros que sí obtuvieron mejoras postest.

Con respecto a las calificaciones escolares, pudimos comprobar, como ya lo hicimos para los grupos completos, que las calificaciones obtenidas por el grupo de mejoradores, sólo correlacionaron positivamente con las puntuaciones obtenidas seis meses antes, es decir, en la fase pretest; mientras que las puntuaciones dinámicas, que reflejaban la valoración de un proceso de aprendizaje, no se relacionaron positivamente con las notas. El análisis de las comparaciones entre los grupos de mejoradores y no mejoradores, nos ofreció un dato revelador con relación al contraste entre las valoraciones realizadas por el profesorado tutor y las calificaciones escolares. En efecto, cuando se realizó el contraste por separado para ambos grupos, pudimos constatar que la correlación entre las calificaciones y la valoración promedio de los tutores, pasaba a ser positiva para los mejoradores, mientras que en el grupo de no mejoradores, el valor de correlación seguía siendo negativo. De este modo, el dato paradójico de correlación negativa entre calificaciones y valoración del tutor, quedaba situado sólo en el grupo de sujetos no mejoradores. Por otra parte, los datos de los análisis de regresión pusieron de manifiesto nuevamente que para el grupo de mejoradores, el porcentaje de varianza explicada por la puntuación pretest se mostraba a todas luces insuficiente para predecir la puntuación

postratamiento, siendo las puntuaciones de mejora aquellas que mejor explicaban los puntajes finales.

En definitiva, en línea con las observaciones y los análisis efectuados por Budoff (1987a, 1987b), hemos constatado la relevancia de los resultados obtenidos por aquellos sujetos que, partiendo de puntuaciones bajas, obtuvieron ganancias postratamiento. De este modo, en palabras de Fernández-Ballesteros y sus colaboradores, la distinción cualitativa realizada, nos otorgaría información sobre el elemento clave de un proceso de ED, es decir, sobre si se obtienen mejoras posteriores a la implementación del tratamiento. Así mismo, hemos constatado la relevancia de las diferencias establecidas entre mejoradores y no mejoradores con relación al nivel de predicción del dispositivo EDPL sobre el rendimiento escolar o sobre las calificaciones escolares; y sobre todo, con relación a las valoraciones efectuadas por el profesorado.

Por lo que respecta al análisis del valor de predicción mostrado por el dispositivo EDPL en la muestra de secundaria, hemos de destacar los altos niveles de correlación que se establecen entre la puntuación dinámica y el conjunto de valoraciones efectuadas por los docentes sobre el rendimiento escolar. Por otra parte, ya expresábamos que la correlación negativa que mantiene esta puntuación con la puntuación de mejora, constituía un resultado, cuanto menos, inesperado, ya que en esta submuestra se produce un incremento significativo de la puntuación postest sobre la inicial y, por otra parte, el valor medio de la puntuación dinámica es muy alto; quizás la explicación esté en que este valor medio de la puntuación dinámica ha estado muy por encima de la puntuación postest, y ha sido ésta última la que sí ha correlacionado muy significativamente con la puntuación de mejora. En este sentido, las valoraciones del profesorado aplicador sobre la actuación de los sujetos en EDPL, reflejadas a nivel cuantitativo en las puntuaciones dinámicas, han superado notablemente a la propia actuación de los sujetos en las pruebas estáticas, lo cual, dado que el dispositivo de evaluación dinámica recoge aspectos y procesos que no se encuentran reflejados en las pruebas estáticas, podría ser indicio de que los resultados en las pruebas criterio no conseguirían recoger una parte importante de la información (Guthke y Beckmann, 2000; Resing, 2000). La correlación positiva de la puntuación dinámica con las valoraciones del profesorado contribuiría a confirmar esta hipótesis. El nivel de correlación positivo entre la puntuación dinámica y las calificaciones escolares en esta etapa, parece estar relacionado con el hecho de que en secundaria hayamos recabado un mayor número de calificaciones numéricas y basadas en la actuación y el rendimiento real del sujeto en el área de Lengua. Esto habría contribuido a que los contrastes con las calificaciones muestren resultados menos llamativos. Por ejemplo, para este grupo, las notas correlacionan positivamente y de forma muy significativa con la valoración del tutor, lo que era inicialmente esperable, ya que, como dijimos cuando analizábamos los resultados de primaria, han sido los mismos profesores los que han realizado la valoración sobre el rendimiento y los que han calificado finalmente el área.

Por otra parte, los datos muestran que la puntuación obtenida en el pretest, mediante la aplicación de la prueba de criterio, no está relacionada positivamente con la puntuación de mejora, ni así mismo con la puntuación dinámica, extraída ésta de la participación de los sujetos en las sesiones de EDPL. Este dato es sumamente importante y esperable desde la perspectiva dinámica, ya que la puntuación inicial, obtenida con una prueba estática, no debía estar especialmente relacionada con una puntuación que expresa la mejora de los sujetos, ni con una puntuación que refleja la actuación de los mismos en

un procedimiento de evaluación-intervención dinámico. En definitiva, las puntuaciones obtenidas de la pasación de las pruebas criterio pretest y postest mantienen entre sí una alta correlación, pero no parecen mostrar capacidad suficiente para explicar o predecir la actuación de los sujetos en el contexto escolar ni, así mismo, en un procedimiento dinámico de evaluación.

Además de la constatación, al igual que ya ocurriera en la muestra de primaria, de que para el grupo de control, las puntuaciones postest y de mejora explicarían la mayor parte de la varianza observada, en contraposición a los datos obtenidos para el grupo experimental, para el que las puntuaciones dinámicas contribuirían de forma significativa a la explicación del rendimiento académico; otro aspecto especialmente relevante y relacionado con la comparación entre los grupos experimentales, es el que tiene que ver con las diferencias observadas entre las valoraciones efectuadas por los docentes sobre el rendimiento escolar y sobre el progreso del alumnado en los grupos experimental y control, así como con relación a las calificaciones finales. Ya apuntábamos que la constatación de diferencias relevantes entre las valoraciones asignadas a ambos grupos, podría ofrecer una valiosa información sobre el rendimiento efectivo de los sujetos; información complementaria, que iría más allá de la que podíamos obtener de los resultados en las pruebas criterio. En este sentido, pudimos comprobar que las valoraciones efectuadas por los tutores sobre el rendimiento escolar, las valoraciones sobre el progreso de los sujetos durante la fase experimental, así como las calificaciones finales otorgadas, fueron muy superiores para los sujetos que formaban parte del grupo experimental, llegando las diferencias a ser estadísticamente muy significativas para el progreso y las calificaciones. El dato de las valoraciones positivas al grupo de tratamiento refuerza notablemente los resultados, ya positivos, obtenidos por este grupo en esta etapa, estableciendo un matiz considerable en cuanto a la ausencia de diferencias significativas directas entre los resultados obtenidos por los grupos experimental y control. Si bien, en algunos de los contrastes, observamos que el grupo de control también obtiene resultados postest positivos, las valoraciones efectuadas por los tutores vendrían a establecer algunas diferencias cualitativas entre las ganancias de ambos grupos, al tiempo que informarían —en el caso de las calificaciones— de la obtención de mejores resultados finales para el grupo de tratamiento en el área académica relacionada con la propuesta, en función de los criterios académicos adoptados.

Los análisis de regresión efectuados nos mostraron, al igual que ya ocurriera con relación a la prueba criterio de primaria, que la puntuación dinámica conseguía explicar y predecir de forma significativa la mayoría de los valores de rendimiento considerados. Tanto para el rendimiento escolar y el progreso de los sujetos, valorados por los tutores, como para estos mismos valores, evaluados por el profesorado de apoyo, la puntuación dinámica resultó ser la variable que mejor lograba explicar las puntuaciones otorgadas al alumnado, obteniendo valores significativos con relación a la bondad de ajuste y siendo además la única variable que era incluida en la ecuación de regresión cuando se realizaron análisis de regresión por pasos. Con relación a la valoración del progreso, pudimos constatar que la inclusión en la ecuación de las puntuaciones pretest, que correlacionaban negativamente con la valoración efectuada, producía un incremento significativo del valor de ajuste del modelo. Este resultado parece lógico, por cuanto una valoración positiva del progreso de los sujetos, se relacionaría bien con una superación de las puntuaciones pretest; en efecto, una correlación negativa entre pretest y progreso, siempre que la valoración de progreso sea positiva (en este caso, el

promedio de 4,26 indica que se han dado más valoraciones positivas que negativas), apuntaría resultados a favor de nuestra hipótesis. Por otra parte, como hemos apuntado anteriormente, la puntuación dinámica no se mostró esencial en la predicción de la puntuación posttest, ni tampoco en la de la puntuación de mejora. En cambio, obtuvo el valor de correlación más alto con las calificaciones finales, aunque no lograba explicar este valor de forma significativa. En síntesis, podríamos concluir que la contribución adicional de la puntuación dinámica, especialmente en lo que respecta a las valoraciones efectuadas por el profesorado sobre el rendimiento y el progreso de los sujetos, resultó especialmente relevante para la explicación de la actuación de los sujetos en esta etapa.

Con relación a los análisis efectuados para los grupos de mejoradores y no mejoradores, queremos hacer referencia a que, en función de algunas de las valoraciones obtenidas por los sujetos que obtuvieron pérdidas postratamiento; valoraciones efectuadas, tanto por los tutores como por el profesorado aplicador, cabe interpretar, como apuntábamos con anterioridad, que los beneficios de la aplicación del dispositivo podrían haberse ampliado al grupo que, sin embargo, no logró ganancias en las pruebas de criterio. Señalábamos como posibilidades para explicar esta inconsistencia, el hecho de que la propia prueba se hubiera mostrado inadecuada para mostrar posibles logros en estos sujetos, o bien la existencia de un posible efecto «techo», ya que sus puntuaciones pretest se encontraban bastante por encima de las obtenidas por el grupo de mejoradores, incluso a las que obtuvieron éstos tras la fase de aplicación.

Los análisis de regresión efectuados para estos subgrupos, mostraron nuevamente la contribución esencial de las puntuaciones dinámicas en la explicación de los distintos valores contemplados, aunque, probablemente debido al menor número de efectivos en los contrastes, fue menor el grado de significatividad estadística de los mismos. Uno de los aspectos más destacables de estos análisis, lo constituye el hecho de que, de forma sistemática, la puntuación de mejora obtuviera una correlación negativa con las valoraciones efectuadas sobre el rendimiento y el progreso de los sujetos del grupo de no mejoradores. En línea con lo que apuntábamos anteriormente, la mayoría de las valoraciones para estos sujetos, incluidas las calificaciones escolares, fueron positivas; incluso algunas se situaron por encima de las que obtuvieron los mejoradores. De este modo, el valor de la puntuación de mejora —negativo en estos casos— no se correspondió con las valoraciones efectuadas por los tutores y los aplicadores del dispositivo. En efecto, las distintas valoraciones efectuadas sobre el rendimiento y el progreso, así como las calificaciones finales otorgadas, habrían puesto de manifiesto una actuación positiva de los no mejoradores; aunque esto no se habría visto reflejado en su ejecución de las pruebas de criterio. Por otra parte, como consecuencia lógica de estos datos, la puntuación dinámica actuó como la variable que mejor predecía las valoraciones asignadas por el profesorado a los sujetos que no obtuvieron ganancias posttest, con valores de correlación muy altos.

Por otra parte, los análisis efectuados sobre el valor de predicción del dispositivo con relación a la escala de ajuste personal-social, revelaron una fuerte conexión entre la puntuación dinámica global y la puntuación dinámica que hacía referencia exclusiva a los procesos de ajuste. Ya comentamos la relevancia de estos datos, no solo con respecto al grado de consistencia del dispositivo dinámico, sino también en referencia a la estrecha relación que hemos encontrado entre el rendimiento general de los sujetos en las sesiones del dispositivo y la puesta en práctica de habilidades relacionadas con los procesos de ajuste. Por otra parte, tanto la puntuación dinámica global, como la

puntuación dinámica de los procesos de ajuste, mostraron un valor de predicción del rendimiento escolar en su conjunto que fue significativamente superior al de las puntuaciones obtenidas en la escala APSL, ya fueran iniciales o postratamiento. Este valor de predicción de las puntuaciones dinámicas fue también puesto de manifiesto con relación a la actitud de los sujetos ante la lectura, valorada por el profesorado tutor y especialista. De esta forma, las puntuaciones dinámicas se mostraron más relacionadas con la actitud de los sujetos relativa a la lectura, que las puntuaciones pretest o postest obtenidas en la escala de ajuste. Estos resultados guardan estrecha conexión con otros obtenidos por nosotros mismos en un estudio reciente (Navarro y Mora, 2003b), en el que un grupo de profesores que aplicaron una versión dinámica de un instrumento de evaluación de metaconocimientos en la lectura, coincidían en observar y valorar un mayor número de actuaciones relacionadas con procesos y elementos de carácter socio-personal y emocional, mediante el proceso dinámico seguido.

Otro aspecto destacable con relación al valor de predicción del dispositivo en el ámbito de los procesos de ajuste, lo constituyen las diferencias encontradas entre las valoraciones otorgadas por el profesorado a los sujetos del grupo experimental y del grupo control. En este sentido, como hemos apuntado anteriormente, pudimos comprobar que todas las valoraciones promedio del grupo control, incluidas las calificaciones académicas finales, se encontraban sensiblemente por debajo de las que fueron dadas a los sujetos del grupo de tratamiento. Por otra parte, encontramos otros datos, relativos al grupo de control, que nos permitían establecer una comparación relevante con los obtenidos por el grupo experimental. En efecto, los valores superiores de correlación pretest, con respecto a los postest, obtenidos por el grupo control con relación al rendimiento, ponían de manifiesto que la valoración inicial se habría mostrado suficiente para informar sobre el rendimiento de los sujetos, dato que también se veía confirmado en parte por el mayor valor de correlación de la puntuación pretest con las calificaciones escolares. En este sentido, apuntábamos que la ausencia de tratamiento habría hecho válida la predicción inicial del pretest sobre el rendimiento. En el grupo experimental, en cambio, la aplicación del dispositivo dinámico habría contribuido de forma significativa a la explicación del rendimiento escolar. Como tuvimos oportunidad de comprobar, para este grupo, tanto la puntuación dinámica global, como la relativa a los procesos de ajuste, obtenían niveles de correlación con el rendimiento escolar valorado por los tutores y por los aplicadores, que se situaban significativamente por encima de los obtenidos por la puntuación pretest.

Los análisis de regresión realizados para el grupo de control mostraron con claridad que, mientras el pretest figuraba como la única variable que lograba explicar de forma significativa el rendimiento escolar y el progreso de los sujetos, valorado por los tutores; la ecuación de regresión del grupo experimental contemplaba únicamente la puntuación dinámica global y la puntuación dinámica de los procesos de ajuste, logrando explicar ambas variables el rendimiento y el progreso de forma significativa. La puntuación dinámica global también figuró como la única que predecía de forma significativa el rendimiento valorado por los aplicadores, mientras que la puntuación postest lo hacía para el grupo de control. En cuanto a la valoración del progreso por parte de los aplicadores, pudimos observar que fue el pretest el que mejor predecía este valor para el grupo experimental, si bien es cierto que las dos puntuaciones dinámicas lograban explicar un porcentaje de varianza similar. Las calificaciones académicas fueron predichas por el postest en el grupo experimental, aunque pudimos constatar que la conjunción de las dos puntuaciones dinámicas, así como también la conjunción de la

puntuación dinámica global y la puntuación de mejora, lograban explicar un porcentaje mayor de varianza que el postest, resultando ambos modelos igualmente significativos estadísticamente. Mientras tanto, los análisis efectuados para el grupo control, mostraron la ausencia de factores significativos que lograran predecir las calificaciones. Con relación a la valoración de la actitud y del progreso observado, mostramos que la puntuación dinámica global fue el único factor determinante en la explicación de las valoraciones efectuadas sobre la actitud ante la lectura, así como sobre el progreso observado. Por otra parte, para la valoración de la actitud efectuada por los aplicadores, fue la puntuación dinámica de los procesos de ajuste, la única variable que lograba explicar significativamente la actitud; mientras que la valoración del progreso volvía a destacar a la puntuación pretest como la variable que mejor predecía este valor, aunque en este caso, el contraste del modelo no resultara significativo. También en esta ocasión, la conjunción de las puntuaciones dinámicas predecía el progreso de forma similar.

En cuanto a la prueba de CI, los análisis sobre el valor predictivo de las distintas puntuaciones, revelaron que las obtenidas en el pretest de Cattell mantenían una correlación muy significativa con el postest de esta misma prueba. En este sentido, se ha argumentado (Mora, 1998) a favor de lo «esperable» de este dato, sosteniendo que de obtenerse un resultado contrario habría que cuestionarse la idoneidad de la composición de las muestras o la práctica concreta de aplicación de la prueba. En efecto, la puntuación CI demuestra no ser fácilmente modificable, y en numerosos estudios la correlación pretest-postest resulta muy significativa estadísticamente, con valores similares a los obtenidos por nosotros. Igualmente, en los estudios para la validación de la prueba, los valores de fiabilidad test-retest son muy elevados (Cattell y Cattell, 1974/2001). Esta «estabilidad» documentada de la prueba criterio utilizada, incrementa así mismo la validez de las posibles diferencias observadas entre las fases inicial y final como fruto de otros factores que no sean la propia prueba. En este sentido, pudimos observar que la puntuación de mejora también correlacionaba de forma positiva y muy significativamente con el postest, incrementando notablemente su valor de correlación con respecto al obtenido con la puntuación pretest. Este dato, unido al de las ganancias observadas en el GE, apunta a que la influencia del proceso de intervención que tuvo lugar en el intervalo entre ambas pasaciones del test, es decir, la aplicación del dispositivo EDPL, habría resultado muy significativa. Una mayor confirmación de estas observaciones la obtuvimos cuando analizamos diferencialmente los datos de los análisis de regresión para el grupo experimental y el grupo de control.

Pudimos comprobar, por otra parte, que tanto la puntuación dinámica de los procesos metacognitivos, como la global, incrementaron sus valores de correlación en la fase postratamiento, con respecto a los que se establecieron con las puntuaciones pretest. Aunque estos valores de correlación no fueran muy altos, el hecho de que se produzca un incremento en el postest, habiéndose obtenido ganancias para el grupo experimental, es indicativo de un mayor ajuste entre las valoraciones realizadas por el profesorado sobre la actuación en el proceso de aplicación dinámica, por una parte, y los resultados obtenidos tras el tratamiento por los sujetos experimentales, por otra. Así mismo, también son superiores los valores de correlación de las puntuaciones dinámicas con la puntuación de mejora en Cattell, que los obtenidos por los puntajes pretest.

Otro aspecto importante lo constituye determinar qué puntuación, estática o dinámica, habría mantenido una relación mayor con el rendimiento escolar. Tradicionalmente se ha entendido que el CI mantiene unos niveles de correlación altos con el rendimiento

escolar; numerosos estudios apoyan esta tesis (Mora, 1998; Sternberg y Grigorenko, 2003; Cattell y Cattell, 1974/2001). Sin embargo, en nuestras hipótesis figuraba que una evaluación dinámica, basada en el proceso de aprendizaje, y contextualizada con relación al contenido de evaluación, sería un factor explicativo más importante que la información aportada por la aplicación estática de una prueba, más aún si ésta es de carácter descontextualizado. En efecto, los resultados obtenidos muestran claramente la superioridad de los niveles de correlación obtenidos por las dos puntuaciones dinámicas consideradas con relación al rendimiento escolar, con respecto a los niveles observados para las puntuaciones CI. Aunque estas últimas incrementan sus niveles de correlación en el postest y obtienen valores significativos con esta puntuación para la valoración del progreso efectuada por los tutores y para la valoración del rendimiento efectuada por los aplicadores, los valores de correlación obtenidos por las puntuaciones dinámicas son estadísticamente muy significativos al nivel 0,01 para las valoraciones del rendimiento efectuadas por los tutores y por los aplicadores, así como para la valoración del progreso efectuada por los tutores. De esta forma, podemos concluir, en nuestro caso, que han sido las puntuaciones obtenidas de la valoración del proceso de aprendizaje las que mejor se han relacionado con la valoración del rendimiento escolar realizada por el profesorado. Estos resultados parecen ir un poco más allá de lo expresado por Resing (2000), cuando plantea que las puntuaciones dinámicas del LIR habrían aportado una validez predictiva adicional sobre la base correlacional de una prueba de inteligencia estática con el rendimiento escolar. Así mismo, con relación a las calificaciones escolares en el área de Lengua, también se obtuvieron valores de correlación más altos para las puntuaciones dinámicas que para las puntuaciones CI, aunque en este caso, los valores no llegaron a ser significativos. En definitiva, la puntuación dinámica ha mostrado una mayor relación con el rendimiento escolar que las puntuaciones estáticas, tanto aquellas obtenidas de pruebas relacionadas con el contenido a evaluar (las pruebas de comprensión y las de ajuste personal-social), como las obtenidas de pruebas de CI.

Los análisis de regresión por pasos para valorar el poder de predicción de las puntuaciones dinámicas sobre el rendimiento escolar, también nos mostraron que la puntuación dinámica global fue el factor que mejor explicaba las valoraciones sobre el rendimiento efectuadas por los tutores y por los aplicadores, así como la valoración del progreso realizada por los tutores. En este sentido, procedimos a valorar el poder de predicción de las puntuaciones dinámicas en función del nivel intelectual de los sujetos. Guthke y Wingenfeld (1992) habían observado en sus estudios que el poder predictivo de las Pruebas de aprendizaje era significativamente mayor para los sujetos con un nivel intelectual más bajo, en comparación con los sujetos de nivel medio y superior. De esta forma, la información que se podía obtener a raíz de la aplicación dinámica, se tornaría especialmente relevante para la población de sujetos que presentaba déficits cognitivos. En nuestro caso, los resultados obtenidos de estos análisis, aunque revelan una cierta tendencia en la dirección apuntada anteriormente, no descartan que el valor de predicción de las puntuaciones dinámicas EDPL, pueda aplicarse también a los sujetos que obtienen puntuaciones medias en la prueba de inteligencia. En cualquier caso, la aplicación experimental del dispositivo de evaluación dinámica, ha tenido lugar sólo en sujetos que fueron seleccionados precisamente por presentar dificultades de aprendizaje relacionadas con la lectura, incluyendo al alumnado con discapacidad cognitiva. De esta forma, el número de efectivos con altas puntuaciones en la prueba de CI a los que haya sido aplicado EDPL, ha sido escaso; con lo que el análisis comparativo y diferencial en función de la variable cognitiva, no presenta excesiva validez. Por otra parte, algunos de los datos que hemos expuesto a lo largo de la investigación, parecen alinearse sin

reservas con las observaciones de Guthke y Wingelfeld; en este sentido, han sido precisamente aquellos sujetos con mayores dificultades de aprendizaje en lectura, así como aquellos que presentaban discapacidad cognitiva, los que han obtenido mejoras más significativas. Esto nos habla, en principio, de que serían estos sujetos los que más beneficios podrían obtener de la aplicación del dispositivo EDPL en el contexto escolar.

Por su parte, la puntuación dinámica de los procesos metacognitivos obtuvo valores similares a los de la puntuación dinámica global para los contrastes relativos al valor de predicción sobre el rendimiento escolar, logrando explicar significativamente el modelo cuando se procedía a introducir de forma individual esta variable en la ecuación. Por otra parte, la puntuación de mejora actuó igualmente como la variable que mejor predecía el progreso de los sujetos valorado por los aplicadores. De este modo, las diversas puntuaciones dinámicas consideradas, incluyendo la puntuación de mejora, fueron las que mostraron un mayor poder de predicción del rendimiento académico y el progreso, valorado por el profesorado. Con relación al grupo de control, los análisis no nos mostraron ningún factor determinante en la explicación del rendimiento, el progreso o las calificaciones escolares.

Con relación a la validez ecológica de las pruebas dinámicas y a su poder de predicción sobre el rendimiento escolar de los sujetos evaluados, pudimos comprobar en algunos de los análisis efectuados sobre los distintos modelos dinámicos, que la capacidad diferencial demostrada por las pruebas dinámicas para anticipar el desempeño de los alumnos en pruebas escolares, no había sido en absoluto concluyente (Sternberg y Grigorenko, 2003). En este sentido, se había apuntado por parte de varios autores de orientación dinámica, la posibilidad de que el rendimiento escolar no fuera un criterio óptimo para evaluar la validez de un modelo dinámico. Las razones estarían relacionadas con las características asignadas a la valoración de este criterio, más cercana a un *producto* que a un *proceso* (Lidz, 1991). Las pruebas convencionales de aptitud estarían en mejores condiciones para predecir el rendimiento escolar, según estos autores, que las pruebas dinámicas. Estas argumentaciones conducen lógicamente a la necesidad de obtener criterios más dinámicos y procesuales que el rendimiento escolar en pruebas académicas, para poder establecer relaciones con los resultados de la aplicación de modelos dinámicos de evaluación. En este sentido, se había descrito que a medida que la correspondencia era mayor entre los procedimientos usados en la aplicación dinámica de las pruebas, y la metodología empleada en la acción educativa, mayor era también el valor predictivo de las pruebas (Sternberg y Grigorenko, 2003).

En línea con las ideas expresadas anteriormente, anotábamos en el capítulo dedicado a la evaluación complementaria, con relación a la validez predictiva de las pruebas de ED, las dificultades para establecer relaciones positivas entre medidas del potencial de aprendizaje obtenidas en situaciones de interacción y de aprendizaje colaborativo, dirigidas en su mayor parte a la evaluación de procesos cognitivos y metacognitivos; y medidas del rendimiento escolar, que podrían expresar valoraciones sobre aspectos o aptitudes más específicas, al tiempo que estar fundamentadas sobre procedimientos de evaluación más estáticos (Lidz, 1991; Kozulin, 2000; Sternberg y Grigorenko, 2003). En este sentido, no sería de extrañar que las puntuaciones dinámicas de EDPL no hayan mantenido niveles altos de correlación con las puntuaciones postest (que podría ser considerada una medida del rendimiento en contenidos del área de Lengua) o con las calificaciones escolares. Por otra parte, como hemos tenido oportunidad de comprobar, las puntuaciones dinámicas han sido las que mejor predecían el rendimiento escolar

valorado por los profesores a través de las plantillas construidas a tal efecto. De este modo, es posible que en la valoración efectuada por los docentes sobre el rendimiento escolar, hayan actuado mecanismos distintos que podrían haber influido positivamente en la obtención de niveles significativos de correlación. De este modo, el hecho de que facilitáramos al profesorado unos ítems de valoración cualitativa, así como los procesos que debían ser valorados (un total de siete, seleccionados de entre los propuestos por la administración educativa y adaptados a los contenidos y objetivos de nuestra investigación; más la valoración sobre el progreso de los sujetos), puede haber contribuido a que las valoraciones efectuadas hayan tenido un carácter más procesual y menos estático, por lo que podría explicar los resultados positivos y significativos de las correlaciones.

En cuanto a los análisis de regresión efectuados con relación a la prueba de CI, estos nos permitieron comprobar que la puntuación de mejora en el grupo de tratamiento lograba predecir de forma muy significativa la puntuación posttest; el porcentaje de varianza explicada por la mejora en el grupo experimental (48%), revelaría la contribución esencial del tratamiento aplicado. Por otra parte, la conjunción de la puntuación de mejora con las puntuaciones dinámicas (metacognitiva y global) explicaba también de forma significativa las puntuaciones obtenidas en el posttest. Estos resultados, como apuntábamos anteriormente, difieren significativamente de los obtenidos para el grupo de control, en el que el peso de la ejecución de la prueba efectuada en la fase pretest, tuvo un peso mucho mayor en las puntuaciones finalmente obtenidas. Este mayor peso del pretest, unido al menor incremento en el valor de predicción cuando se introducía en el análisis la puntuación de mejora, mostraban una mayor relevancia en el grupo experimental de la puntuación de mejora, es decir, aquella que es reflejo del cambio.

Por otra parte, es destacable que las puntuaciones CI no hayan obtenido correlaciones altas con el resto de pruebas criterio, especialmente con las pruebas de comprensión. Era en cierto modo esperable, a tenor de la mayoría de los estudios, en los que se relaciona de forma significativa la comprensión lectora y el nivel de inteligencia, que estos valores fueran medios o altos con estas pruebas; sin embargo, salvo para la puntuación pretest de ECO 1 y la posttest de ECOS, el resto de correlaciones obtiene valores medios o bajos. En este sentido, las puntuaciones pretest y de mejora obtenidas en las pruebas de comprensión correlacionan significativamente más alto con el posttest que las medidas de inteligencia general, ya sean éstas iniciales o finales. A pesar de esto, y de que no se hayan obtenido valores de correlación muy significativos, parece innegable la aportación que la puntuación CI hace a los niveles de comprensión observados.

Discusión de los resultados del análisis de los cuestionarios EIDAL y EIDAL-MP

El análisis que efectuamos a raíz de las respuestas otorgadas por los docentes y las familias a los cuestionarios sobre expectativas e ideas relacionadas con la atención educativa a las dificultades de aprendizaje de la lectura, nos ofreció algunos datos relevantes con relación a una mejor contextualización de los resultados de nuestro estudio. Se establecieron algunas relaciones significativas entre los elementos recogidos en los cuestionarios y aquellos relacionados con los procesos y contextos de aprendizaje que se hicieron presentes durante la aplicación experimental del dispositivo de evaluación dinámica. El análisis exhaustivo presentado nos permite acceder, al menos

en parte, a los pensamientos, las valoraciones y las actuaciones del profesorado y de las familias participantes, lo que tiene para nosotros una relevancia notable, dado que de este análisis podrían extraerse algunas conclusiones relacionadas con la necesidad de formación del profesorado en el ámbito de la intervención y la evaluación de la lectura en sujetos con dificultades de aprendizaje, así como algunas orientaciones que podrían resultar válidas en este sentido. Nos parece de interés, por tanto, exponer a continuación algunos elementos para la reflexión y la discusión, relacionados con el contenido de estos cuestionarios.

Con relación al cuestionario EIDAL, y a la relevancia de los distintos procesos en la explicación de las dificultades de lectura, pudimos comprobar en análisis anteriores la notabilidad de los procesos metacognitivos y de los procesos de integración texto-conocimientos previos en la explicación de las dificultades de aprendizaje observadas en el alumnado participante del estudio, así como en la explicación de los resultados obtenidos. Estos procesos destacaron por su mayor dificultad, tanto en las pruebas de carácter estático, como en el análisis cualitativo realizado durante el proceso de aplicación del dispositivo experimental. Los resultados del análisis de las respuestas otorgadas por los docentes, nos muestran, sin embargo, que las ideas que manifiesta el profesorado sobre qué dificultades concretas de aprendizaje presentan los alumnos, colocan estas dificultades en aspectos relacionados con procesos que no supusieron especialmente una gran dificultad para los sujetos. La importancia de esta falta de conexión entre los resultados empíricos con relación a los procesos implicados en la lectura que supondrían una mayor dificultad para el alumnado con problemas, y las ideas del profesorado sobre dónde se encuentran mayormente los problemas, podría estar relacionada con la realización de notables esfuerzos dirigidos a la mejora de la lectura, por parte del profesorado, sin que ello vaya acompañado de forma significativa por cambios positivos y duraderos. Por otra parte, la idea reiterada de que el origen de las dificultades de aprendizaje de la lectura se encuentra en el propio alumnado, podría relacionarse, aún de forma no totalmente consciente, con una concepción de las dificultades como algo intrínseco, estable y, por lo tanto, poco modificable. Este aspecto, como tuvimos oportunidad de comprobar a lo largo de la primera parte de nuestro estudio, constituiría un «duro adversario» a los intentos de promover y desarrollar procesos de evaluación dinámica.

Por otra parte, se hizo referencia por parte de los docentes a que la ausencia de motivación con relación a la lectura, se debería esencialmente a factores centrados en el propio alumno, según las opiniones expresadas en el cuestionario; mientras que el contexto escolar figuraba como factor poco determinante en la configuración de los procesos motivacionales. El análisis de las respuestas otorgadas a la cuestión relativa a los elementos que serían necesarios para incrementar la motivación en el alumnado, sin embargo, nos mostraba una mayoría de referencias relativas al proceso de enseñanza-aprendizaje, y concretamente a las estrategias metodológicas que sería idóneo poner en práctica. En este sentido, es especialmente relevante para el profesorado, la búsqueda de textos atractivos para el alumnado y que conecten con sus intereses; así mismo, se destacó la importancia de la presentación atractiva de los textos por parte de los docentes.

Al igual que la valoración cualitativa de las respuestas dadas por los docentes, el análisis de las correlaciones entre las puntuaciones de mejora y las respuestas en EIDAL nos permitió establecer, aún en modo exploratorio, algunas relaciones significativas

entre la mejora obtenida por los sujetos en las distintas pruebas criterio contempladas, y determinadas expectativas e ideas expresadas por el profesorado en los cuestionarios. Por ejemplo, con relación al bloque que recoge las respuestas sobre el origen de las dificultades de aprendizaje de la lectura y los factores que influyen en éstas, es destacable que, mientras que se observan correlaciones positivas y significativas entre la mejora en comprensión y algunas respuestas que centran las dificultades en el propio alumno (dificultades de atención y memoria, procesos de asociación G-F, dificultades en la comprensión, habilidades metacognitivas, etc.); se establece una correlación negativa entre la puntuación de mejora y las respuestas que indican que las dificultades se situarían en la ausencia de motivación, implicación o dedicación en el contexto familiar. Este dato sería especialmente relevante, por cuanto podría revelar la ausencia de cambios positivos precisamente en aquellos casos en los que el profesorado ha destacado las situaciones de deprivación socio-cultural y familiar como principal factor explicativo de las dificultades relacionadas con la lectura. Entre las posibles explicaciones a este hecho, podría estar una cierta delegación, por parte del profesorado, de la responsabilidad del cambio educativo en el ámbito socio-familiar; o dicho de otro modo, las ideas que sitúan la explicación de las dificultades en variables que se encuentran fuera del ámbito de influencia directa del profesorado, podrían estar actuando como fuerza generadora de expectativas negativas hacia las posibilidades de mejora del alumnado, lo que podría a su vez explicar un bajo rendimiento. Por otra parte, una explicación que quizás podría ajustarse más a la realidad de lo acontecido en nuestro estudio, destaca la relevancia real de las dificultades encontradas por los docentes que han desarrollado su trabajo en centros educativos situados en zonas de actuación educativa preferente. En este sentido, ya hemos hecho referencia a los problemas relacionados con el absentismo de parte del alumnado participante de estos centros.

Algo similar a lo que hemos tratado de explicar en las líneas anteriores, ha ocurrido cuando las respuestas del profesorado se han dirigido a situar las dificultades en el contexto escolar, especialmente en el proceso de enseñanza-aprendizaje. Hallamos algunas correlaciones negativas entre este tipo de respuestas y la mejora de los sujetos. Explorando la procedencia de estas respuestas, encontramos que fueron dadas desde un centro educativo situado en una zona deprimida; la mayoría del alumnado participante en el estudio desde este centro, como tuvimos oportunidad de comprobar cuando analizamos los resultados en las distintas pruebas criterio, obtuvo malos resultados en las pruebas de comprensión y no consiguió mejoras significativas. Ya hemos analizado también algunas de las circunstancias especiales que caracterizaron el proceso de aplicación del dispositivo en el alumnado que formó parte del grupo experimental en este centro, por lo que algunas de las respuestas dadas por la profesora que cumplimentó el cuestionario han obtenido de forma reiterada una correlación negativa con las puntuaciones de mejora.

Otro aspecto de interés, relacionado en esta ocasión con las dificultades de aprendizaje de la lectura concretadas en un alumno del aula, tiene que ver con la correlación negativa que mantiene la mejora en comprensión con algunas respuestas que sitúan las dificultades en procesos de integración textual y en algunos procesos metacognitivos. En este sentido, una posible interpretación de estos resultados podría ser que cuando se señala que éstas últimas son las dificultades principales que presenta un alumno, no se habrían producido mejoras en las pruebas criterio de comprensión; sin embargo, hemos podido comprobar, sobre todo para el subgrupo de secundaria, que los sujetos que

obtuvieron ganancias postratamiento, incrementaron significativamente sus puntuaciones en las dimensiones metacognitivas de las pruebas. Por otra parte, es importante reseñar la existencia de respuestas que destacan la dificultad de un mismo tipo de procesos, obteniendo en cambio niveles de correlación contrapuestos con la puntuación de mejora; estos resultados se deben a que la mejora fue desigual en los distintos grupos de aplicación y, por consiguiente, también los resultados del análisis correlacional. Lógicamente, no en todos los casos, las valoraciones, actuaciones, ideas o expectativas expuestas en el cuestionario, se han correspondido con similares resultados en las pruebas criterio de nuestro estudio.

A lo largo del análisis efectuado, hemos tratado de establecer algunas conexiones entre las respuestas otorgadas por los docentes a distintas cuestiones. En este sentido, con relación al *como* se realizaba la evaluación de la comprensión, y a la información que se cree relevante recabar en esta evaluación, encontramos una mayor correspondencia entre las respuestas que hacían referencia a aspectos relacionados con los procesos de asociación G-F y con los procesos de integración textual; en cambio, observamos cierta falta de correspondencia entre la relevancia que otorgan algunos docentes a las respuestas que indican la necesidad de recabar información sobre aspectos relacionados con los procesos metacognitivos y de ajuste personal-social, y la ausencia de referencias a estos aspectos dentro de las acciones encaminadas a la evaluación de la comprensión.

Con relación al cuestionario EIDAL-MP, hemos expuesto igualmente un análisis minucioso de las respuestas proporcionadas por las familias en respuesta a las cuestiones recogidas en el protocolo entregado. Este análisis nos ha permitido obtener información adicional a la que se deriva de la valoración del trabajo desarrollado en las aulas. De esta forma, hemos podido acceder a las ideas, expectativas y valoraciones realizadas por las madres y los padres del alumnado participante, y establecer relaciones entre éstas y el proceso de aprendizaje puesto en práctica durante la aplicación de EDPL.

El análisis de aquellos aspectos en los que el profesorado y las familias mostraron acuerdo en sus expectativas e ideas, o, por el contrario, expresaron pensamientos divergentes en sus respuestas, constituye un aspecto relevante de nuestra revisión del contenido de ambos cuestionarios. En este sentido, la percepción de las familias coincidía con la de los profesores en señalar los aspectos relativos al propio alumnado como aquellos que explicarían mejor las dificultades de aprendizaje de la lectura. Las familias, al igual que los docentes, hicieron referencia especialmente a los procesos de ajuste personal-social, así como, en segundo lugar, a los procesos psicológicos subyacentes. Con relación a los factores de índole social-familiar y escolar, coincidieron también con los profesores en señalar en menor medida su relevancia. Entre los factores citados, también figuraban la motivación y el interés que debían despertar en los alumnos los textos que éstos debían leer. Por otra parte, también mostraron acuerdo con los docentes en señalar, como factor social, la «competencia» que suponía, con relación a la lectura, la incidencia de los medios audiovisuales.

En cuanto a los análisis correlacionales entre las puntuaciones de mejora y las respuestas expresadas por las familias, expusimos en nuestra revisión algunos datos que, sin duda, son relevantes con relación a posibles acciones educativas encaminadas a la optimización de los procesos educativos, tanto en el ámbito escolar como en el familiar. En este sentido, algunas de las respuestas relativas al contexto escolar como factor

explicativo de las dificultades de aprendizaje, mantienen una relación positiva con las puntuaciones de mejora en la escala de ajuste personal-social. El atractivo de los textos y de las actividades relacionadas con la lectura que debe trabajar el alumnado, supone uno de los elementos más destacados por las familias, ya sea como factor explicativo de las dificultades, o como aspecto fundamental a trabajar para incrementar la motivación y el interés de los sujetos.

Discusión sobre la incidencia de la aplicación del dispositivo de evaluación dinámica EDPL en la mejora del proceso de evaluación

La presencia en un estudio como el que presentamos, de efectos imprevistos o complementarios a los esperados, es sumamente habitual (Mora, 1998). De este modo, se han observado y descrito, por parte de los aplicadores, efectos de la aplicación del dispositivo EDPL, que han sido reflejados en sus observaciones y valoraciones, sin que, por el contrario, éstos se hayan mostrado en los resultados de las pruebas de comprensión y de ajuste. Algunos de estos efectos de la aplicación del dispositivo, relativos a la mejora del componente *control-metacognitivo* —componente que, por otra parte, podría estar detrás de las ganancias observadas en CI y en las demás pruebas—, guardan cierto paralelismo con los resultados observados y expuestos en la valoración complementaria del programa de enriquecimiento cognitivo Comprender y Transformar (Mora, 1998). En este caso, se alude a la relevancia que pudieron tener con relación a las ganancias observadas, aspectos como la reducción de la impulsividad, la mejora en la planificación de estrategias o la mayor atención prestada en las tareas propuestas; así como otros aspectos más relacionados con procesos de ajuste personal-social, como una imagen más positiva de sí mismo y de sus posibilidades para poder solucionar las actividades, o la mejora en la motivación hacia las tareas. En este sentido, nosotros hemos expuesto algunas observaciones similares efectuadas por los aplicadores. Estas observaciones han quedado reflejadas y codificadas en las *Hojas de registro y evaluación*, mediante la valoración de los indicadores propuestos.

En línea con lo apuntado anteriormente, hemos visto la relevancia de los procesos metacognitivos y de los procesos de ajuste personal-social en la explicación del proceso de comprensión y de las dificultades encontradas. Los datos provenientes de las pruebas criterio, como, sobre todo, los que se desprenden del análisis de la aplicación del dispositivo EDPL, apuntan al hecho de que estos procesos, junto con el que hace referencia a la integración del texto con los conocimientos previos, tienen una importancia central en la configuración final del proceso de lectura y comprensión, tanto para el alumnado de primaria como para el de secundaria. En este sentido, las altas correlaciones de las puntuaciones dinámicas de los procesos metacognitivos y de ajuste, con la puntuación dinámica global, así lo confirma. Así mismo, en función de los promedios obtenidos, ambos procesos parecen encontrarse entre los que representan una mayor dificultad con relación al aprendizaje de la lectura, especialmente para el alumnado que ha formado parte de los grupos experimental y control en nuestro estudio. Por otra parte, en las sesiones del Grupo de Trabajo, los aplicadores confirmaron de forma sistemática, a lo largo de la fase experimental, la relevancia de estos procesos en el análisis de las dificultades presentes en los sujetos. Estos datos, y de forma particular aquellos en los que se destaca la relevancia de los procesos metacognitivos como fuente de especiales dificultades relacionadas con el aprendizaje y dominio de la lectura, confirmarían, a su vez, la quinta de nuestras hipótesis formuladas.

Con relación a la posible incidencia de la aplicación experimental del dispositivo EDPL en el proceso de enseñanza y evaluación de la comprensión, se ha destacado por parte de los integrantes del Grupo de Trabajo la constatación de que el trabajo desarrollado habría aportado estrategias nuevas y relacionadas con el afrontamiento de la evaluación y la intervención en las dificultades de aprendizaje relacionadas con la lectura; así mismo, se hace alusión a una mayor capacidad de observación y análisis de aspectos relativos a la evaluación de procesos de comprensión, que antes no eran observados. En la misma línea se pronunciaron en referencia a la valoración global del proceso y de los resultados obtenidos con la aplicación de EDPL; los aplicadores consideraron especialmente positivo el propio aprendizaje relacionado con las estrategias de evaluación e intervención; así como con la teoría sobre el proceso de comprensión de textos y sobre los diversos procesos implicados. También destacaron en sus análisis los beneficios que habría tenido la aplicación del dispositivo en el alumnado participante.

Discusión sobre las propuestas de mejora del dispositivo EDPL

La valoración realizada por los expertos sobre el contenido de las pruebas de comprensión y ajuste; así como la aplicación piloto de las mismas en la fase inicial de nuestra investigación, nos ofrecieron la posibilidad de analizar, y modificar en su caso, el contenido y la estructura de estas pruebas, elaboradas con el propósito de establecer unos criterios estándar para la medición pretest y postest del rendimiento de los sujetos. Por otra parte, hemos analizado igualmente algunos aspectos mejorables del dispositivo EDPL, o que requerirían de ulteriores investigaciones específicas. De este modo, las valoraciones efectuadas por los integrantes del GT durante el proceso de aplicación, sobre los indicadores de evaluación, que se materializaron posteriormente en las puntuaciones dinámicas; las valoraciones de estos mismos profesionales, efectuadas posteriormente sobre el dispositivo EDPL; y, por último, los propios resultados obtenidos en las diversas pruebas criterio, han sido las fuentes, además de nuestra propia observación, en las que nos hemos basado para determinar en qué aspectos podría modificarse EDPL para incrementar su potencia. En este sentido, a través del cuestionario de valoración, los aplicadores hicieron mención de mejoras relacionadas con el contenido de las actividades (presentación más atractiva de textos y actividades; adecuación del contenido a los diferentes niveles o etapas a las que podría dirigirse; y dificultad y amplitud de las tareas propuestas). También se especificaron mejoras relacionadas con el proceso de evaluación, que hacían referencia sobre todo a aspectos de organización del material (presencia de los indicadores de evaluación propuestos, en el *Cuaderno de Hojas de registro y evaluación*; contar con más espacio para realizar las anotaciones de la mediación; tener la posibilidad de recoger la información de forma colectiva en aplicaciones grupales; etc.); aunque también se recogieron otros elementos en este apartado, que implicarían modificaciones a otro nivel y en los que nos detendremos a continuación. Por último, se recogía una sugerencia relativa a la mejora de la teoría de base del dispositivo. Esta idea expresaba la necesidad de explicitar más detalladamente los microprocesos implicados en cada una de las tareas propuestas en el dispositivo.

Con relación a las posibles mejoras en el proceso mismo de evaluación, queremos detenernos en dos aspectos que consideramos de especial relevancia para nuestra propuesta; uno era referido a la posible normalización del proceso de mediación, y el otro a la posibilidad de contar con versiones similares de una misma actividad con el objetivo de poder evaluar la evolución en la práctica de las estrategias trabajadas. Con

relación al primero de estos aspectos citados, ya expusimos en el capítulo dedicado al análisis de nuestro dispositivo dinámico, que en cada una de las actividades presentadas, se recogía una propuesta de mediación y evaluación de los procesos metacognitivos. En este apartado de la tarea, quedaba expresada nuestra propuesta de aplicación de los patrones metodológicos de evaluación dinámica, con el objetivo de facilitar la obtención de información valiosa para una posterior intervención en las dificultades de lectura. Como hemos apuntado posteriormente, estos patrones metodológicos han sido revisados, complementados y organizados a la luz de la aplicación experimental de EDPL, como fruto de la labor desarrollada por los profesionales que realizaron la aplicación. Así mismo, en el citado capítulo, apuntábamos una opción decidida por centrarnos en la determinación —a través de la aplicación dinámica del dispositivo— de aquellos procesos en los que el sujeto presenta dificultades, así como de aquellas condiciones y estrategias que optimizarían el proceso de enseñanza y aprendizaje. En este sentido, la normalización del proceso de mediación, o en su defecto una mayor sistematización de la propuesta, podría contribuir a homogeneizar el proceso de evaluación propuesto, con indudables ventajas como, por ejemplo, mayores posibilidades de establecer comparaciones entre los resultados obtenidos por diferentes aplicadores. Por otra parte, una de las características definitorias y esenciales de nuestra propuesta, es la flexibilidad en la aplicación del dispositivo, así como la posibilidad de contextualizar incluso los instrumentos de evaluación. De este modo, pensamos que una normalización, en sentido estricto, del proceso de mediación, podría entrar en contradicción con la flexibilidad que creemos necesaria aplicar para obtener el máximo rendimiento de una propuesta como la nuestra, más enfocada a la mejora de los procesos de lectura en el contexto escolar, que al establecimiento de medidas estándar susceptibles de comparación. En sintonía con la reflexión de los autores del método de las sugerencias graduadas (Campione y Brown, 1987; Campione, 1989); como, así mismo, con lo que ya apuntábamos en nuestra propia introducción de este trabajo, los procedimientos de evaluación dinámica pueden encontrarse en una disyuntiva difícil de solventar. Por un lado, se hace necesaria la búsqueda de diseños y modelos teóricos y metodológicos que aporten suficiente validez y fiabilidad, con el fin de generalizar sus aplicaciones; por otro, existe la constatación de que un énfasis excesivo en la cuantificación y normalización de los procedimientos dinámicos, implica una cierta desnaturalización del proceso dinámico de evaluación, con la consiguiente renuncia a la posibilidad de la obtención de información significativa.

Con relación al segundo de los elementos que citábamos, es decir, la posibilidad de contar con versiones similares de una misma actividad, ya hemos discutido la relevancia de la contextualización de los instrumentos de evaluación en nuestra propuesta original. A pesar de esto, somos conscientes de que la mayoría de los profesores que se adentraran en la aplicación de una propuesta como la que presentamos, agradecerían la presencia de textos o actividades que permitieran evaluar la evolución de las estrategias que son objeto de trabajo.

11. Conclusiones

Una vez expuestos en los capítulos de Resultados, los efectos obtenidos en la presente investigación, así como el proceso de evaluación seguido para valorar la aplicación experimental de nuestro dispositivo de evaluación dinámica; hemos procedido, en la primera parte de este bloque, a la discusión e interpretación de los diversos datos generados. Nos resta, por tanto, sintetizar a continuación una serie de conclusiones con el doble objetivo de ser útiles a los efectos de abrir futuras vías de análisis, así como de ser tenidas en cuenta en nuevas propuestas de intervención y evaluación encaminadas a la optimización de los procesos de enseñanza y aprendizaje de la lectura, especialmente con relación al alumnado que presenta dificultades de aprendizaje. Las conclusiones que presentamos guardan, lógicamente, relación con los objetivos planteados, así como también con las hipótesis de trabajo que quedaron formuladas al inicio de la segunda parte de este estudio.

Con relación a la comprensión de textos

1. La evaluación dinámica de procesos implicados en la lectura, mediante la aplicación del dispositivo EDPL, ha mostrado su eficacia en la mejora significativa del nivel de comprensión para el alumnado con dificultades en su lectura. En este sentido, con relación a la etapa de primaria, los análisis efectuados en función del tercero de los criterios adoptados para el análisis, nos mostraron que el grupo experimental obtenía mejoras significativas postratamiento. Estas mejoras se producían para la puntuación total y para la dimensión *Metaconocimiento estratégico*, hallándose igualmente diferencias notables entre las puntuaciones pretest y postest para la mayoría de las dimensiones de la prueba. Por otra parte, el grupo de control no obtuvo ganancias estadísticamente significativas bajo este tercer criterio. De este modo, las ganancias del grupo experimental tenían lugar precisamente en aquellos sujetos con mayores dificultades de lectura y que, a su vez, presentaban bajas puntuaciones CI.

2. El dispositivo EDPL se habría mostrado igualmente eficaz para lograr mejoras significativas en la comprensión de textos para aquellos sujetos que obtuvieron puntuaciones iniciales bajas en la prueba. En este sentido, con relación al segundo de los criterios de análisis adoptado, encontramos ganancias significativas postratamiento a favor del grupo experimental. Estas ganancias fueron para la puntuación total y también para la dimensión metacognitiva *Metaconocimiento estratégico*. Los contrastes también mostraron diferencias nítidas entre las fases pretest y postest para la casi totalidad de las dimensiones contempladas en la prueba. De este modo, el dispositivo también habría mostrado un efecto positivo y significativo favorable a nuestra primera hipótesis, aunque la obtención de ganancias muy significativas en el grupo de control, restaría significación a los resultados del grupo de tratamiento.

3. Con relación al criterio de designación previa del profesorado sobre la presencia de dificultades de aprendizaje; es decir, teniendo en cuenta al grupo completo de sujetos que conformaron los grupos experimental y control de la investigación, debemos admitir que el dispositivo de evaluación dinámica EDPL no se ha mostrado eficaz para lograr un incremento significativo postratamiento en el grupo experimental; por el contrario, el grupo de control sí obtuvo diferencias significativas entre las fases pretest y postest del estudio. En el capítulo de Discusión hemos abordado en profundidad las posibles razones que podrían estar detrás de estos resultados. En este sentido, las ganancias significativas halladas para el GC1, podrían revelar una descompensación

relevante en la selección y composición de la muestra de control cuando son tenidos en cuenta todos los sujetos. Por otra parte, los resultados de análisis posteriores, nos mostraron que la mayor presencia en el grupo de tratamiento de sujetos que recibían apoyo a la integración y de sujetos con deprivación socio-cultural, habría tenido una influencia significativa en la obtención de bajas puntuaciones por parte de este grupo. Así mismo, los análisis efectuados en función de los diversos subgrupos contemplados en la muestra, nos mostraron que las ganancias observadas por el GC1 sólo se mantuvieron para el subgrupo de sujetos que no precisaba de atención educativa específica por parte del profesorado de apoyo a la integración o audición y lenguaje. En este sentido, las comparaciones realizadas tomando en consideración el primero de los criterios de análisis, podrían adolecer de validez suficiente relativa a la selección inicial de los sujetos que fueron adscritos a los grupos de la investigación.

4. El análisis diferencial de los resultados obtenidos en esta etapa en función de los grupos de aplicación, nos mostró diferencias relevantes. En este sentido, el factor *aplicador*, unido al factor *contexto de aplicación*, habría influido decisivamente en las puntuaciones globales del grupo de tratamiento en la prueba de comprensión de primaria. Los análisis efectuados con carácter exploratorio, con aquellos grupos de aplicación en los que las circunstancias contextuales no fueron especialmente desfavorables, dieron como resultado que el grupo experimental obtuvo ganancias postratamiento cercanas a la significatividad estadística.

5. Los análisis efectuados para el reducido grupo de sujetos que, en primaria, respondía al perfil de *disléxicos disfonéticos o fonológicos*, en función de sus puntuaciones en las pruebas de comprensión y CI, mostraron ganancias notables en la fase postest. Los sujetos del GE obtuvieron una mejora promedio de 5 puntos, aunque esta diferencia no resultó estadísticamente significativa. Por su parte, los sujetos del GC1 que respondían a este perfil también obtuvieron mejoras postest importantes.

6. Con relación a la comprensión de textos en la etapa de secundaria, podemos destacar que el dispositivo de evaluación dinámica EDPL, ha mostrado también un efecto significativo favorable a nuestra primera hipótesis experimental. En este sentido, los resultados muestran ganancias muy significativas para el grupo de tratamiento en los tres criterios de análisis adoptados. Con relación al criterio que aúna la designación del profesorado sobre la presencia de dificultades de aprendizaje, y la obtención de bajas puntuaciones en comprensión y en la prueba de CI; los resultados muestran ganancias significativas para el GE en la puntuación total y en dos de las dimensiones de la prueba: *Metaconocimiento estratégico* y *Conocimientos previos*. Por otra parte, las ganancias observadas inicialmente en el GC1 no se mantuvieron cuando se re-analizaron los resultados eliminando de los contrastes a aquellos sujetos que no recibían atención educativa específica.

7. Con relación al segundo de los criterios formulados, se hallaron nuevamente diferencias muy significativas para el grupo experimental a favor de la hipótesis. Estas ganancias fueron obtenidas en la puntuación global de la prueba, así como en tres de las dimensiones contempladas: *Metaconocimiento estratégico*; *Conocimientos previos*; e *Integración texto-conocimientos*. Por su parte, el GC1 también obtuvo mejoras significativas en la puntuación total y en una de las dimensiones, aunque éstas fueron de menor intensidad.

8. En cuanto a los resultados en función del primer criterio de análisis adoptado, es decir, el que tomaba en consideración sólo la designación por parte del profesorado colaborador sobre las dificultades de aprendizaje del alumnado; los resultados nos mostraron ganancias muy significativas a favor del grupo que recibió el tratamiento, tanto en la puntuación global de la prueba como en cinco de las seis dimensiones contempladas: *Concepción sobre la lectura*; *Metaconocimiento estratégico*; *Conocimientos previos*; *Integración texto-conocimientos previos*; y *Supervisión y autoevaluación del proceso de comprensión*. Los efectos positivos de la aplicación dinámica se habrían distribuido entre los efectivos de la muestra, con más del 90% de la evolución de los puntos a favor de nuestra hipótesis experimental. Aunque el GC1 también obtuvo ganancias significativas, los análisis comparativos entre las puntuaciones de mejora de los grupos experimental y control, nos mostraron diferencias muy relevantes a favor del grupo de tratamiento en todas las dimensiones, llegando a la significatividad estadística en la dimensión metacognitiva *Supervisión y autoevaluación del proceso de comprensión*. Estas diferencias se hallaron igualmente para la submuestra formada por sujetos que obtuvieron bajas puntuaciones iniciales en la prueba ECOS, como para la submuestra que obtuvo puntajes iniciales medios o altos.

9. Encontramos diferencias significativas y muy significativas en los contrastes efectuados para la prueba ECOS en función de los subgrupos contemplados en la investigación. Estas diferencias se dieron tanto en el GE como en el GC1, aunque fueron de mayor intensidad para el grupo de tratamiento. Los subgrupos en los que hubo ganancias significativas postratamiento, fueron: el grupo de sujetos con DA en ausencia de discapacidad o deprivación; el grupo de sujetos con discapacidad; el grupo que recibía apoyo a la integración; y el grupo de sujetos que no recibía apoyo, ni presentaba discapacidad o deprivación.

10. Las ganancias significativas obtenidas en secundaria por el grupo experimental, se encontraron también cuando se realizaron los contrastes en función de los distintos grupos de aplicación. A pesar del reducido número de efectivos en algunos de los grupos, las mejoras observadas fueron estadísticamente significativas para la mayoría de ellos, produciéndose ganancias posttest en todos. Las ganancias observadas se produjeron en la puntuación total de la prueba, y también en aquellas dimensiones que habían representado una mayor dificultad para los sujetos: *Conocimientos previos*; *Integración texto-conocimientos*; y *Supervisión y Autoevaluación del proceso de comprensión*. Estos análisis constituyen un indicador relevante con relación a la posible generalización de los resultados obtenidos en esta etapa.

11. Los sujetos del GE que respondieron en la etapa de secundaria al perfil de *disléxicos disfonéticos*, obtuvieron en este caso ganancias significativas y muy significativas, tanto para la puntuación global de la prueba, como para las dimensiones: *Metaconocimiento estratégico*; *Conocimientos previos*; e *Integración texto-conocimientos previos*. Por su parte, los sujetos del GC1 que respondían a este perfil no obtuvieron diferencias pretest-posttest que fueran relevantes.

12. Las valoraciones efectuadas por el profesorado tutor sobre el rendimiento escolar y sobre el progreso experimentado por el alumnado de secundaria durante la fase experimental, así como las calificaciones académicas otorgadas en el área de Lengua, fueron muy superiores para aquellos sujetos que formaban parte del GE, siendo estadísticamente significativas estas diferencias con relación al progreso y a las calificaciones. De este modo, el GE habría obtenido mejores calificaciones finales en el

área de conocimientos directamente relacionada con la investigación. Estos datos vendrían a confirmar y a reforzar los resultados positivos del GE en esta etapa, estableciendo diferencias de carácter cualitativo entre las ganancias obtenidas por ambos grupos.

Con relación al ajuste personal-social

13. La evaluación dinámica de procesos implicados en la lectura, mediante la aplicación del dispositivo EDPL, ha mostrado su eficacia en el logro de ganancias significativas en los procesos de ajuste personal-social. Estas ganancias se produjeron en función del tercero de los criterios de análisis adoptados; es decir, cuando fueron seleccionados los sujetos con bajas puntuaciones iniciales en la escala de ajuste y en la prueba de CI. Las ganancias se produjeron en el GE para la puntuación global de la escala, así como para las dimensiones D1, D9 y D8 y 9. Por su parte, el GC1 obtuvo diferencias importantes a favor del postest, aunque éstas sólo fueron significativas para una de las dimensiones. De este modo, la aplicación de EDPL ha mostrado efectos significativos favorables a la segunda de nuestras hipótesis formuladas.

14. La evaluación dinámica de procesos implicados en la lectura, ha logrado igualmente mostrar su eficacia para producir mejoras muy significativas en los procesos de ajuste personal-social para el GE, cuando se ha adoptado el segundo de los criterios de análisis. Estas mejoras se han producido en la puntuación global de la escala, así como en algunas de las dimensiones contempladas; entre ellas, la dimensión metacognitiva D8 y 9 y la D1. Por su parte, el GC1 no consiguió mejoras significativas para ninguna de las puntuaciones.

15. La aplicación del dispositivo EDPL se ha mostrado eficaz en el logro de mejoras postratamiento en la escala de ajuste para aquellos sujetos que, en primaria, obtuvieron bajas puntuaciones en comprensión. Este subgrupo obtuvo ganancias por encima de los dos puntos y medio para la puntuación global en la escala, que era superior para algunas de las dimensiones contempladas. El grupo de control, por su parte, no obtuvo incremento postest. El GE en secundaria logró una mejora promedio cercana a los 5 puntos, y una mejora estadísticamente significativa para la dimensión D3; mientras que el GC1 obtuvo pérdidas notables en el postest. Por otra parte, la aplicación del dispositivo EDPL también ha ocasionado mejoras postratamiento muy relevantes al subgrupo de sujetos con bajas puntuaciones iniciales, tanto en comprensión como en la escala de ajuste. De esta forma, y con las limitaciones que impone lo reducido de esta submuestra, el dispositivo habría mostrado un efecto significativo sobre los procesos de ajuste personal-social, precisamente, en aquellos sujetos con mayores dificultades en comprensión, y también en los procesos de ajuste.

16. Con relación al criterio de designación previa sobre la presencia de dificultades de aprendizaje; y tomando de nuevo en consideración al grupo completo, incluidos los sujetos de primaria y secundaria, debemos admitir que la aplicación experimental del dispositivo dinámico, no ha logrado mejoras estadísticamente significativas en los procesos de ajuste personal-social para el GE. Sí son destacables, en función del contenido de nuestra segunda hipótesis, las mejoras postratamiento obtenidas en la puntuación global de la escala por el grupo de tratamiento, así como en algunas de las dimensiones contempladas. Por su parte, el GC1 obtuvo pérdidas postest, tanto en la puntuación global, como en el resto de dimensiones.

17. La aplicación del dispositivo EDPL se mostró eficaz con relación al logro de mejoras postratamiento en la escala de ajuste cuando se realizaron análisis en función de la etapa. En este sentido, mientras que los sujetos del GE en primaria sólo obtuvieron ligeras mejoras postratamiento, los contrastes efectuados para el alumnado de secundaria, mostraron ganancias muy significativas para algunas de las dimensiones contempladas en la escala, especialmente para las dimensiones metacognitivas. Por su parte, el GC1 obtuvo pérdidas generalizadas, tanto en primaria como en secundaria. Estos resultados nos permiten delimitar con mayor precisión el perfil de los sujetos que habrían obtenido mayores beneficios de la aplicación del dispositivo y, especialmente, del trabajo desarrollado con relación a los procesos de ajuste. Así mismo, los resultados también nos han revelado que el proceso de mediación desarrollado habría tenido un efecto más potente sobre las dimensiones metacognitivas de la escala. Por último, los efectos observados nos permiten deducir la relevancia de los procesos metacognitivos en el dispositivo dinámico puesto en práctica, así como en el rendimiento observado en lectura; ya que las ganancias obtenidas en las dimensiones metacognitivas de la escala de ajuste, han tenido lugar en la etapa donde se dieron igualmente las mayores ganancias postratamiento en las pruebas de comprensión.

18. Por su parte, los resultados obtenidos en primaria revelan un «igualamiento» entre chicos y chicas con especiales dificultades de aprendizaje de la lectura, con relación al nivel de comprensión de textos y de adaptación personal-social. En secundaria, las chicas han mostrado de forma general mayores niveles de ajuste personal-social, evaluado con la escala APSL. Este mayor nivel de las chicas llega a ser estadísticamente significativo para la dimensión que hace referencia a la *percepción del apoyo que ofrece el contexto escolar y familiar, y de las condiciones en las que se desenvuelve*.

19. La aplicación del dispositivo EDPL ha mostrado su eficacia para producir mejoras significativas en los procesos de ajuste personal-social para el subgrupo de sujetos con dificultades de aprendizaje de la lectura, en ausencia de discapacidad o deprivación. Las ganancias significativas se han producido en la puntuación global de la escala y en la dimensión metacognitiva D8 y 9. En tanto, el GC1 experimentaba ligeras pérdidas postest. En este sentido, los contrastes entre las puntuaciones de mejora de ambos grupos, establecían diferencias nítidas en la puntuación global y en el resto de dimensiones, llegando a ser significativas en la D1. Así mismo, el subgrupo de sujetos del GE que no recibía apoyo específico, ni presentaba discapacidad o deprivación socio-cultural, obtuvo ganancias postratamiento, aunque éstas no llegaron a ser significativas. Estas ganancias se produjeron de forma generalizada, en la puntuación global y en las dimensiones de la escala, obteniendo en algunas de ellas mejoras por encima de los 8 puntos. De nuevo fueron las dimensiones metacognitivas, las que obtuvieron mejores resultados, acercándose a la significatividad estadística. Por su parte, el GC1 obtuvo pérdidas postest generalizadas. Con relación a los subgrupos formados por los sujetos con deprivación socio-cultural, y por aquellos que recibían apoyo a la integración, la aplicación del dispositivo dinámico ha originado ligeras ganancias postratamiento en los procesos de ajuste para el GE. Estas ganancias no llegaron a ser significativas estadísticamente. Sólo los valores relacionados con la dimensión metacognitiva D9 para el subgrupo *recibe apoyo*, se acercaron a la significatividad. En cuanto al GC1, los resultados para ambos subgrupos muestran igualmente pequeñas ganancias postest. Por otra parte, con relación al subgrupo de sujetos con discapacidad, la aplicación del dispositivo EDPL no ha logrado producir mejoras postratamiento en los niveles de ajuste para el GE. Los resultados muestran ligeras pérdidas postest, tanto en la

puntuación global, como en las dimensiones de la escala. Por su parte, el GC1 también obtuvo pérdidas.

20. Los análisis efectuados con relación a los distintos grupos de aplicación, nos revelaron que la aplicación del dispositivo EDPL ha originado, para la mayoría de los grupos de aplicación, incrementos notables postratamiento en algunas de las dimensiones de la escala. Las mayores diferencias y el mayor número de ganancias en los distintos grupos, se produjeron en las dimensiones metacognitivas de la escala, así como en la dimensión D1.

21. En definitiva, la evaluación dinámica de procesos implicados en la lectura, a través de la aplicación del dispositivo EDPL, ha tenido una repercusión significativa sobre los procesos y contenidos relacionados con el conocimiento metacognitivo, incluyendo así mismo, el metaconocimiento sobre estrategias de comprensión (Mora Merchán y Mora, 2000; Mateos, 2001). Ha sido especialmente este metaconocimiento estratégico (dimensiones D8; D9; y D8 y 9 de la escala), aquél sobre el que la aplicación del dispositivo EDPL habría actuado, originando cambios significativos generalizados; mientras que los procesos de ajuste personal-social contenidos en el resto de dimensiones, habrían experimentado una transformación más limitada, o más difícilmente mensurable. En este sentido, como apuntábamos anteriormente, las valoraciones promedio que recibieron los sujetos del GE, incluidas las calificaciones finales, fueron notablemente superiores a las que recibió el GC1. De igual forma, también la valoración efectuada sobre la actitud de los sujetos con relación a la lectura, estableció diferencias relevantes entre las valoraciones realizadas en ambos grupos, a favor del GE.

22. En el transcurso de la aplicación experimental del dispositivo EDPL, se produjo una evolución positiva en el grado de ajuste personal-social con relación a la lectura, para el GE de primaria; y, en menor medida, para el GE en secundaria. La evolución pretest-postest de las correlaciones entre las puntuaciones obtenidas en la escala de ajuste personal-social y en la prueba de comprensión ECO 1, fue positiva para el GE en los tres criterios de análisis adoptados. Los resultados muestran un mayor ajuste entre las puntuaciones globales de ambas pruebas, una vez aplicado el dispositivo; y para una de las dimensiones (D2), la convergencia con la puntuación global en comprensión, llega a ser significativa. Los sujetos que obtuvieron mejoras postratamiento en las dimensiones metacognitivas de la escala de ajuste, puntuaron más alto en el postest de la prueba de comprensión. En el GC1, la evolución de estas correlaciones fue marcadamente negativa. Por su parte, la evolución pretest-postest de las correlaciones en secundaria, fue positiva en función de los dos primeros criterios de análisis, aunque marcadamente negativa en función del tercero. Los resultados de los análisis mostraron que los cambios positivos producidos, tanto en la prueba de comprensión ECOS, como en la escala de ajuste, no se produjeron en la misma dirección, observándose débiles correlaciones entre las puntuaciones obtenidas en las dos pruebas. Para el GC1 en esta etapa, la evolución tuvo un carácter ligeramente más positivo.

23. La convergencia entre las mejoras obtenidas en las dimensiones metacognitivas de la escala de ajuste, y las obtenidas en la prueba de comprensión ECO 1, avala una mayor fiabilidad de las mejoras observadas. Sin embargo, las conclusiones finales sobre este aspecto deben reflejar también las débiles correlaciones obtenidas entre las puntuaciones obtenidas en las dos pruebas, así como el exiguo número de sujetos que obtuviera ganancias tanto en ajuste como en comprensión. Con relación a la etapa de

secundaria, es destacable que un número importante de sujetos obtiene ganancias en ambas pruebas, aunque los análisis correlacionales entre las puntuaciones o entre las diferencias promedio de las dos pruebas, no permiten extraer conclusiones definitivas sobre la fiabilidad de las mejoras. Las ganancias significativas obtenidas por el GE en la etapa de secundaria, no se han visto acompañadas proporcionalmente por ganancias en la escala de ajuste personal-social.

Con relación a la inteligencia general

24. La evaluación dinámica de procesos implicados en la lectura, mediante la aplicación del dispositivo EDPL, ha mostrado su eficacia en el logro de ganancias cognitivas muy significativas. Estas ganancias se han producido en función de los distintos criterios adoptados para el análisis, siendo mayores cuando los sujetos presentaban puntuaciones iniciales más bajas, tanto en comprensión como en la misma prueba de inteligencia. Las ganancias tuvieron lugar tanto en primaria como en secundaria, y también se extendieron al análisis realizado en función de los distintos subgrupos considerados, aunque sólo fueron muy significativas estadísticamente para el subgrupo de sujetos con discapacidad. Por su parte, el GC1 no obtenía mejoras posttest. En este sentido, los resultados obtenidos en los diversos contrastes, verifican sustancialmente nuestra tercera hipótesis experimental.

25. El proceso de mediación efectuado habría tenido una relevancia fundamental en la explicación de las mejoras significativas observadas en la prueba de inteligencia. Los análisis efectuados con los *mejoradores* reforzaron significativamente la relevancia de los patrones de mediación en las ganancias observadas. La intervención preferente y específica sobre los procesos metacognitivos durante la aplicación experimental del dispositivo, a través de la puesta en práctica de los patrones metodológicos de mediación; habría influido significativamente sobre el mayor rendimiento de los sujetos en la prueba de inteligencia en la fase posttest. Esta influencia se habría producido como consecuencia de un efecto de *transferencia* de aquellas habilidades y estrategias generales, trabajadas sobre procesos implicados en el dominio específico de la comprensión de textos, e interiorizadas en el transcurso de su participación en las sesiones de aplicación del dispositivo EDPL.

26. En definitiva, los resultados parecen confirmar que la aplicación del dispositivo de evaluación dinámica EDPL habría producido un efecto positivo significativamente mayor en la población de sujetos con dificultades cognitivas y de aprendizaje más acusadas, y que partían de puntuaciones más bajas en las pruebas de criterio. En este sentido, en línea con lo reflejado en numerosos estudios revisados en nuestro trabajo (Budoff, 1987a, 1987b; Campione y Brown, 1987; Guthke y Beckmann, 2000; Carlson y Wiedl, 1992b, 2000; Swanson, 2000), la aplicación de EDPL ha originado mejoras muy significativas en los sujetos que presentan mayores dificultades, tanto en funciones y dominios específicos, como en inteligencia general. De este modo, los resultados obtenidos habrían configurado empíricamente la población que obtendría mayores beneficios de la aplicación del dispositivo en el contexto escolar.

Con relación a la mejora del proceso de evaluación y al proceso de aplicación

27. Los profesionales que realizaron la aplicación experimental del dispositivo de evaluación dinámica EDPL, han expresado en las sesiones del GT, así como en los cuestionarios de valoración del dispositivo, un alto grado de satisfacción con relación a

la idoneidad del mismo para la obtención de información valiosa, relevante y orientada a la mejora del proceso de enseñanza y aprendizaje de la lectura. Así mismo, han expresado —individual y colectivamente— en los análisis cualitativos realizados durante las sesiones del GT, un alto grado de satisfacción con relación a las sesiones de formación y supervisión de la puesta en práctica de EDPL. Los aplicadores destacaron especialmente el propio aprendizaje realizado con relación a la estrategia metodológica de evaluación dinámica, así como con los aspectos teóricos relativos al proceso de comprensión de textos y a los diversos procesos implicados en la lectura. De este modo, la aplicación del dispositivo EDPL habría aportado a los aplicadores nuevas estrategias de evaluación e intervención en las dificultades de aprendizaje relacionadas con la lectura; así mismo, los aplicadores habrían incrementado su capacidad de observación y análisis de procesos y habilidades relacionadas con la lectura, enfocando su atención y sus valoraciones hacia indicadores de evaluación considerados relevantes para la obtención de información valiosa, que previamente no eran observados. El análisis conjunto del proceso de aplicación experimental, llevado a cabo por los profesores que participaron del GT, así como el análisis posterior del proceso y de los resultados obtenidos, ha mostrado con claridad que la propuesta de evaluación dinámica contenida en el dispositivo EDPL, puede ser aprendida e implementada en el contexto escolar por parte de los docentes. En este sentido, el proceso de formación inicial y de supervisión de los aplicadores, se ha revelado como un factor *clave* en el éxito de la implementación de la propuesta. Futuras aplicaciones del dispositivo EDPL en el contexto escolar, deben ir acompañadas de procesos de formación inicial y de supervisión, a través de fórmulas de trabajo en equipo, y guiados por profesionales con experiencia y conocimientos sobre la metodología y los contenidos de la ED y del dispositivo EDPL.

28. Los elementos del dispositivo EDPL considerados *clave* por los aplicadores con relación a la obtención de información valiosa para una posterior intervención, tienen que ver con la propia propuesta metodológica. En este sentido, los aplicadores destacaron: (a) el procedimiento de evaluación dinámica seguido; (b) la mediación de procesos cognitivos y metacognitivos; y (c) la posibilidad de centrarse en los indicadores de evaluación propuestos y en los patrones metodológicos de mediación, durante la aplicación. Por otra parte, también destacaron otros elementos que a su juicio fueron relevantes; éstos estaban relacionados con el registro de la información; los diversos procesos contemplados en la propuesta, destacando los procesos metacognitivos; el trabajo colectivo de análisis y discusión en el GT; y la propuesta concreta de actividades. Con relación al proceso de aplicación, los elementos considerados *clave*, a juicio de los aplicadores, en la explicación de las mejoras observadas, habrían tenido que ver, por una parte, con: (a) la propuesta metodológica realizada; es decir: con la mediación de procesos cognitivos y metacognitivos implicados en la lectura; la lentificación del proceso en aquellos bloques de actividades donde se detectaron dificultades; o la puesta en práctica de las actividades del dispositivo de forma lúdica y participativa. Por otra parte, también habrían estado relacionados con (b) los procesos trabajados durante la aplicación; es decir: las estrategias de aprendizaje trabajadas en la mediación; el trabajo sistemático sobre las habilidades metacognitivas; o la intervención sobre los diversos procesos, en función del bloque de actividades trabajado. Por último, los aplicadores también destacaron (c) el esquema de actividades propuesto y la estructura interna de las tareas, como aspectos relevantes en la explicación del cambio.

29. Entre las principales dificultades que hemos encontrado en la aplicación experimental del dispositivo EDPL, se encuentran: (a) la difícil conjugación en la

práctica, de una aplicación contextualizada y flexible de la propuesta, con los requerimientos psicométricos de validez y fiabilidad; (b) los problemas de carácter institucional-organizativo que han tenido lugar en algunos centros, y a los que ya hemos hecho referencia en páginas anteriores. Estos problemas habrían dificultado una implementación óptima del dispositivo dinámico; y (c) la dificultad para trasladar la dinámica del trabajo propuesto en las sesiones de aplicación, al trabajo sobre lectura y comprensión en el aula ordinaria de los alumnos. Con relación a la primera de estas dificultades, queremos destacar que la aplicación en el contexto escolar del dispositivo EDPL, con los objetivos y el método contenidos en la propuesta original, y sin los requerimientos psicométricos de una aplicación experimental, atenuaría o eliminaría el problema que mencionamos. En cambio, el segundo y el tercero de los problemas expuestos podrían ser, al menos en parte, consustanciales a la puesta en práctica, en el contexto escolar, de una propuesta como la que presentamos. Sin perjuicio de que, en futuras aplicaciones, nuestra actuación tratara de atenuar variables de tipo institucional, o contemplara actuaciones encaminadas a la coordinación y cooperación en el trabajo de análisis e intervención con los sujetos evaluados (aspecto ya contemplado y desarrollado en nuestra propuesta original); entra dentro de lo posible que nos encontremos con dificultades similares.

30. Las circunstancias y condiciones contextuales en las que se han desarrollado algunas de las aplicaciones del dispositivo, han actuado de modo conservador con relación a las hipótesis experimentales, reduciendo de forma significativa, en algunos casos, las condiciones óptimas de aplicación en las que debería haberse evaluado la aplicación experimental del dispositivo EDPL. Algunas de estas circunstancias habrían podido tener un peso significativo en el desarrollo y en los resultados obtenidos por alguno de los grupos. En este sentido, debemos destacar que la obtención de mejoras significativas en las distintas pruebas de criterio; los valores de predicción significativos del dispositivo con relación al rendimiento escolar y a las puntuaciones de mejora de las pruebas; así como las valoraciones especialmente positivas del profesorado que realizó la aplicación experimental, sobre las ganancias notables que habrían experimentado algunos sujetos en el transcurso de la aplicación; todo ello bajo circunstancias que distaban de ser favorables, reforzaría significativamente la verificación de las hipótesis formuladas, al tiempo que contribuye a confirmar la relevancia y la necesidad de implementar procedimientos dinámicos de evaluación de los procesos implicados en la lectura, en el contexto escolar.

31. A pesar de que no ha sido posible la puesta en práctica de algunas de las acciones contempladas en la propuesta original, encaminadas a una mayor contextualización del proceso de evaluación, así como a un mayor conocimiento del contexto educativo y social donde se desarrollaba el proceso; hemos logrado recoger información relevante relacionada con los diversos ámbitos en los que se desarrolla el estudio. De este modo, se ha obtenido información a través del profesorado, acerca de aquellas circunstancias o condiciones socio-culturales que podrían influir en el proceso o en los resultados; incluyendo esta información en los análisis efectuados. Así mismo, hemos obtenido información acerca del contexto educativo y socio-familiar a través del análisis de los cuestionarios EIDAL y EIDAL-MP.

32. El análisis de las expectativas e ideas del profesorado y de las familias acerca de las dificultades de aprendizaje de la lectura, así como de las correlaciones entre las puntuaciones de mejora de los sujetos y las respuestas otorgadas por docentes y madres/padres en los cuestionarios; nos ha revelado datos significativos con relación a

una mejor comprensión y contextualización del proceso y de los resultados obtenidos en nuestra investigación. Se han establecido algunas relaciones significativas entre los aspectos recogidos en las diversas respuestas facilitadas en los cuestionarios, y los procesos y contextos de aprendizaje que han tenido lugar en la fase experimental del estudio. De los análisis efectuados con base en los cuestionarios EIDAL y EIDAL-MP, así como de los realizados por los aplicadores en su valoración de la aplicación experimental del dispositivo, se desprende la relevancia y la necesidad de la implementación de programas de formación del profesorado relacionados con los procesos implicados en la comprensión de textos, con su evaluación en el contexto escolar, así como con la intervención educativa más eficaz encaminada a la mejora de los sujetos. Así mismo, de los análisis llevados a cabo pueden deducirse orientaciones que podrían resultar valiosas en este sentido.

33. Las mejoras observadas en los procesos de comprensión y de ajuste personal-social, han ido más allá de los resultados obtenidos de los contrastes con las pruebas criterio. En primer lugar, los aplicadores han informado de mejoras muy significativas en las habilidades trabajadas en el dispositivo; mejoras que no siempre se han visto reflejadas en las puntuaciones obtenidas por los sujetos en las pruebas estáticas. En este sentido, los sujetos del GE recibieron valoraciones relativas al rendimiento escolar y al progreso experimentado durante el tiempo de aplicación experimental, sensiblemente mayores que las obtenidas por los sujetos del GC1. Por otra parte, los resultados promedio obtenidos por el GE en una situación de aprendizaje mediado, como es la que ha tenido lugar durante la aplicación del dispositivo EDPL, han sido sensiblemente mejores que los que este grupo ha podido obtener en una prueba de comprensión de textos o de ajuste personal-social, administrada en forma convencional. De este modo, la valoración cuantificada de un proceso de aprendizaje, ha obtenido valores superiores a los que se han obtenido de la valoración del producto de la comprensión. Por último, ha tenido lugar en algunos casos una evolución positiva de la correlación establecida entre las puntuaciones dinámicas y el postest, con respecto a la mantenida con el pretest. El proceso de evaluación dinámica y de aprendizaje mediado, desarrollado en la aplicación dinámica efectuada, ha ofrecido de este modo mayores posibilidades de evaluación y de intervención en la *zona de desarrollo potencial* de los sujetos, y habría tenido lugar sobre procesos *en desarrollo*. En este sentido, las ganancias significativas obtenidas por el GE, habrían confirmado la hipótesis relativa a una optimización del trabajo autónomo tras un periodo de aprendizaje mediado (Vygotski, 1934/1995; Rogoff, 1993; Werstch, 1988; Luque, Ortega y Cubero, 1997; Navarro, 2004).

34. La valoración promedio del rendimiento escolar efectuada por los tutores y por los profesores de apoyo, fue sensiblemente mayor para los *mejoradores* que para el grupo de *no mejoradores*; de esta forma, las valoraciones sobre el rendimiento, realizadas por los tutores, estarían en consonancia con las valoraciones sobre el proceso de aprendizaje llevado a cabo durante la aplicación del dispositivo, efectuadas por los aplicadores. Así mismo, estos resultados contribuyen a validar el procedimiento utilizado para evaluar el rendimiento académico, ya que tutores y apoyo habrían efectuado sus valoraciones en consonancia con los resultados obtenidos por los sujetos experimentales en las pruebas criterio. Con relación a las puntuaciones dinámicas, las valoraciones promedio entre *mejoradores* y *no mejoradores* también son significativamente favorables a los primeros.

35. En general, hemos encontrado mayor sintonía entre valoraciones efectuadas sobre procesos, y con un carácter más cualitativo; por ejemplo, entre: (a) las valoraciones del

proceso de aplicación dinámica; (b) la valoración del rendimiento escolar y del progreso del alumnado; y (c) las opiniones del profesorado aplicador. De forma más puntual, hemos encontrado confluencia de resultados entre algunos valores de las pruebas criterio y las puntuaciones dinámicas o la valoración del rendimiento escolar.

36. Las ganancias postratamiento obtenidas por los sujetos del grupo experimental, pueden ser consideradas como muy relevantes, en función de la distancia existente entre el contenido y la metodología de las tareas que han debido realizar los sujetos en las sesiones del dispositivo EDPL, y el de las pruebas de criterio pretest y postest. De este modo, las ganancias no expresarían sólo diferencias pretest-postest, sino que mostrarían la puesta en práctica de procesos de transferencia «lejana» (Campione y Brown, 1987; Campione, 1989). Por otra parte, el análisis de los resultados que muestran las mejoras postratamiento, nos ofrece información sobre la apropiación e integración de la información por parte de los sujetos, al tiempo que sobre aquellos procesos que han sido objeto de mediación y han debido ser puestos en práctica en el transcurso de las actividades del dispositivo. Así mismo, los resultados de los análisis sobre el proceso de mediación, y especialmente del análisis de los patrones metodológicos de evaluación dinámica utilizados en la mediación, nos otorgan información acerca de aquellas estrategias metodológicas de interacción docente, que se han revelado determinantes en la obtención de mejores resultados en la resolución de las actividades propuestas; al tiempo que nos estarían ofreciendo información sobre la propensión de los sujetos a beneficiarse de la mediación puesta en práctica por los aplicadores.

37. Hemos extraído del proceso de aplicación experimental del dispositivo EDPL, una serie de patrones metodológicos de mediación que se mostraron eficaces en la mejora de los procesos evaluados, o bien en la obtención de información valiosa para la optimización del proceso de enseñanza-aprendizaje. Estos patrones metodológicos, utilizados en la mediación por los aplicadores, han sido finalmente 19, y aglutinaban a su vez, 178 unidades de información que reflejaron las conductas concretas de mediación puestas en práctica por los aplicadores durante las sesiones del dispositivo dinámico. La matriz construida se estructuró en tres grandes apartados: (a) *aspectos generales*, que contemplaba fundamentalmente patrones de mediación relativos a aspectos generales de la intervención; (b) *mediación del conocimiento metacognitivo y del uso de estrategias*; y (c) *mediación de procesos de ajuste personal-social*. El análisis de estos patrones de mediación ha dado lugar a la determinación de aquellos que habrían sido *claves* en la explicación de las mejoras observadas en las distintas pruebas. Es destacable la relevancia de patrones de mediación pertenecientes a las tres categorías especificadas, así como la mayor significación de su puesta en práctica en aquellos sujetos con mayores dificultades, o con relación a los procesos que mayores dificultades representaron. En este sentido, las mejoras se habrían relacionado fundamentalmente con conductas docentes relativas a la *profundización en la respuesta del alumno o en su justificación, haciendo nuevas preguntas o cuestionando sus respuestas; la recapitulación conjunta del trabajo realizado; la mediación de la discusión y reflexión sobre las estrategias de comprensión; la repetición de una actividad con el objetivo de evaluar la puesta en práctica de un determinado proceso o de ampliar la información; o el apoyo en la comprensión*. Por otra parte, se habrían relacionado con pautas docentes más específicamente dirigidas a la mediación de los procesos implicados en la comprensión, y especialmente de los procesos metacognitivos: *mediación de la conciencia sobre las propias estrategias de comprensión; mediación de la conciencia de la estructura fonológica, léxica y sintáctica; mediación de la conciencia sobre elementos estructurales; mediación de la*

conciencia de las propias dificultades; mediación de la conciencia de lo que el sujeto sabe acerca del contenido de los textos trabajados; mediación del uso de estrategias metacognitivas de planificación y autoevaluación; mediación del uso de estrategias de planificación; mediación del uso de estrategias de supervisión; mediación de estrategias de autoevaluación; mediación de estrategias de transferencia y generalización; mediación de estrategias de integración textual; mediación de estrategias de integración del texto con conocimientos previos. Por último, también se relacionaron de forma relevante con conductas docentes que se relacionaron con los procesos de ajuste: *mediación de la motivación, la actitud y el interés.*

38. Estos patrones metodológicos de ED utilizados en la mediación habrían tenido una relevancia notable en la consecución de las ganancias observadas en aquellos sujetos que obtuvieron mejoras postratamiento en las pruebas de comprensión y de ajuste, partiendo de puntuaciones bajas, es decir, en los *mejoradores*. Los valores más altos de correlación entre patrones de mediación y puntuación dinámica para este grupo, se produjeron especialmente para aquellas puntuaciones relacionadas con los procesos metacognitivos, los procesos de ajuste personal-social, y los procesos de integración texto-conocimientos. De este modo, el proceso de mediación efectuado por los aplicadores, a través de la puesta en práctica de los patrones metodológicos de ED, habría tenido una relevancia especialmente significativa en la obtención de las mejoras observadas en aquellos sujetos con mayores dificultades, o con relación a aquellos procesos que representaron mayores dificultades para todos los sujetos.

39. De forma general, podemos destacar que los diversos análisis psicométricos efectuados para evaluar la validez y la fiabilidad de las pruebas de comprensión ECO 1 y ECOS, así como de la escala de ajuste personal-social APSL, nos ofrecieron índices más que aceptables. Estas pruebas fueron elaboradas por nosotros con el objetivo de servir como criterio para valorar los posibles cambios producidos como consecuencia de la aplicación del tratamiento. Así mismo, los análisis de validación, efectuados previamente a la aplicación de las pruebas por el grupo de expertos, con relación a la adecuación de las pruebas para evaluar la comprensión de textos y los niveles de ajuste personal-social con relación a la lectura; mostraron índices de valoración muy positiva. Con relación a la prueba de comprensión ECO 1 existen, sin embargo, datos relevantes que apuntan a la posibilidad de que esta prueba habría representado un nivel excesivo de dificultad para el alumnado con problemas de aprendizaje. En este sentido, los índices de dificultad obtenidos en los análisis psicométricos, unidos a la valoración efectuada por el grupo de expertos con relación a la excesiva dificultad que podría representar la prueba para el alumnado con dificultades, avalarían esta hipótesis.

40. La estimación realizada sobre la fiabilidad inter-observadores, mediante el cálculo del coeficiente de correlación intra-clase, nos permitió comprobar que la consistencia interna de las observaciones y valoraciones efectuadas por los aplicadores, ha sido especialmente alta, con valores cercanos al acuerdo absoluto con relación a las observaciones realizadas con el total de indicadores de evaluación valorados en las actividades puestas en práctica. Así mismo, la estimación de la fiabilidad con relación a las valoraciones de los indicadores metacognitivos, tuvo valores cercanos al acuerdo absoluto entre los aplicadores.

41. Los análisis diferenciales realizados en función de los grupos de aplicación, muestran ganancias significativas generalizadas cuando se toman en consideración diversos aspectos: la prueba de criterio utilizada; el criterio de selección utilizado para

efectuar los análisis; o los diversos subgrupos considerados con relación a las características cognitivas o socio-culturales. De este modo, bajo la consideración de estos elementos, nuestras hipótesis experimentales habrían sido contrastadas y verificadas en condiciones muy diversas de calidad de aplicación. Así mismo, el factor *aplicador*, unido al factor *contexto de aplicación*, habría tenido una relevancia notable en la explicación diferencial de las mejoras cuando los análisis han sido efectuados con los grupos completos, es decir, con mayor independencia de la toma en consideración de los aspectos mencionados. Por otra parte, en relación al establecimiento de diferencias en los procesos de mediación efectuada por los distintos aplicadores, los análisis muestran diferencias relevantes en la cantidad de anotaciones relativas a la mediación efectuada, aunque estas diferencias no tuvieron su reflejo en las puntuaciones dinámicas globales obtenidas en los diversos grupos.

42. El dispositivo de evaluación dinámica EDPL ha mostrado una consistencia interna notable con relación a su contenido. En este sentido, los análisis efectuados muestran un alto número de correlaciones significativas y muy significativas entre las puntuaciones dinámicas de los distintos procesos contemplados, así como entre éstas y la puntuación dinámica global. Los distintos procesos del dispositivo conformarían un conjunto interrelacionado de elementos que contribuirían a la valoración del proceso de comprensión. También ha mostrado consistencia con relación a la teoría de base en la que está basado, observándose una relación muy significativa entre la puntuación dinámica global y la puntuación dinámica metacognitiva, lo que a su vez refuerza la relevancia de los procesos metacognitivos en la valoración global del proceso de comprensión, al tiempo que la relevancia de la puntuación dinámica metacognitiva en la puntuación dinámica global obtenida en EDPL. El dispositivo, de este modo, habría mostrado también su consistencia y validez en cuanto al sistema de codificación y valoración utilizado.

43. De forma general, la aplicación experimental del dispositivo EDPL, ha mostrado la eficacia e idoneidad de los distintos elementos que conforman la propuesta: objetivos, contenidos, metodología y evaluación. Por otra parte, el análisis efectuado por los propios aplicadores durante la fase experimental, unido a nuestras propias observaciones, apunta la necesidad de introducir ciertos cambios con el fin de optimizar su implementación en el contexto escolar. Estos cambios afectan fundamentalmente al (a) contenido de las actividades propuestas (presentación más atractiva de textos y actividades, y adecuación del contenido a los diferentes niveles o etapas a las que podría dirigirse; y posibilidad de contar con versiones similares de una misma actividad para evaluar la evolución de las estrategias trabajadas y las mejoras del alumnado); así como a la (b) organización del material en el proceso de evaluación (presencia de los indicadores de evaluación en el Cuaderno de *Hojas de registro y evaluación*; contar con más espacio para realizar las anotaciones de la mediación; y posibilidad de anotaciones colectivas en aplicaciones grupales). Por otra parte, el conjunto de patrones metodológicos de evaluación dinámica, derivado del proceso de revisión, complementación y organización final de los patrones utilizados en la mediación, y efectuado a la luz de la aplicación experimental del dispositivo, constituirá la base metodológica fundamental de nuestra propuesta en la fase de mediación.

Con relación a los procesos metacognitivos, su relevancia y su dificultad

44. Los procesos metacognitivos tendrían una relevancia notable en la explicación del proceso de comprensión y de las dificultades encontradas. Los datos cruzados de los

diversos análisis efectuados, apuntan la importancia fundamental de la metacognición en la configuración final del proceso de comprensión, tanto en el alumnado de primaria como en el de secundaria. Los resultados obtenidos muestran la relevancia e idoneidad de actuar especialmente sobre los procesos metacognitivos implicados en la lectura, con el objetivo de mejorar el rendimiento escolar y cognitivo en aquellos sujetos que muestran, precisamente, un bajo funcionamiento metacognitivo.

45. Los procesos metacognitivos y los procesos de integración texto-conocimientos, habrían representado las mayores dificultades para el alumnado. Estos procesos destacaron por su dificultad en las pruebas estáticas de comprensión, así como en las valoraciones realizadas por los aplicadores durante las sesiones del dispositivo EDPL. De este modo, los procesos metacognitivos se encontrarían, para un gran número de sujetos, en la base de las dificultades de aprendizaje relacionadas con la lectura y el proceso de comprensión. La convergencia entre el grado de dificultad que representaron las dimensiones metacognitivas de las pruebas de comprensión, y el que supusieron las habilidades metacognitivas en EDPL, sitúan estos procesos entre los que representaron mayores problemas para el alumnado; lo que contribuye a la validación de nuestra quinta hipótesis. Por otra parte, el análisis de las mejoras con relación a la convergencia inicial observada, otorga resultados positivos para el grupo experimental en secundaria. Este grupo experimenta mejoras significativas en sus puntuaciones metacognitivas de las pruebas criterio de comprensión, logrando que estas dimensiones desaparezcan tras la aplicación del tratamiento, de entre las que representan una mayor dificultad.

46. El trabajo sistemático y específicamente dirigido a la mejora de las habilidades metacognitivas implicadas en la comprensión de textos, se encontraría en la base de la evolución positiva de los resultados obtenidos en las dimensiones metacognitivas de la prueba criterio en secundaria. Las ganancias posttest en las dimensiones metacognitivas, no tuvieron lugar en el GC1, en el que tampoco se produjo una evolución positiva en el grado de dificultad que la dimensión metacognitiva de la prueba representaba para el alumnado.

47. El análisis del proceso y de los resultados obtenidos, nos muestra la posibilidad de establecer vínculos teóricos y empíricos —en el sentido propuesto por Carlson y Wiedl (2000)— entre las *estructuras cognitivas generales* y los *dominios y funciones específicos*. En efecto, la puesta en práctica de habilidades metacognitivas de autorregulación, en el transcurso de las sesiones del dispositivo EDPL, se habría relacionado con un mayor rendimiento en dominios específicos como la comprensión de textos, así como en otras funciones cognitivas específicas. Estos vínculos han sido especialmente puestos de manifiesto en los análisis cualitativos realizados por los profesores aplicadores durante las sesiones de valoración conjunta en el GT, así como en el cuestionario de evaluación del dispositivo dinámico y de su puesta en práctica. Así mismo, a un nivel más cuantitativo, estos vínculos se han puesto de relieve en los contrastes que establecían ganancias significativas entre las puntuaciones iniciales y postratamiento de nuestro estudio, para las pruebas de comprensión; así como en los análisis que relacionaban significativamente las puntuaciones dinámicas metacognitivas con el rendimiento escolar y el progreso de los sujetos.

48. Con relación a la escala APSL, los valores de correlación entre la puntuación global y la puntuación obtenida en las dimensiones metacognitivas, son muy significativos, revelando la importancia de estas dimensiones en los resultados globales de la escala.

Así mismo, estos resultados nos muestran las relaciones entre los procesos de ajuste personal-social, y la metacognición (Mora, 1997; Mora Merchán y Mora, 2000).

Con relación a la aplicación piloto

49. El análisis de la evolución de las puntuaciones en comprensión de textos para el grupo de aplicación piloto, tanto en la etapa de primaria como en la de secundaria, muestra ganancias muy significativas entre las fases pretest y postest. En este sentido, hemos verificado nuestra sexta hipótesis experimental, relativa al progreso del alumnado sin especiales dificultades en los procesos de lectura, en ausencia de intervención específica.

50. Por su parte, el análisis de la evolución pretest-postest en las puntuaciones obtenidas en la escala de ajuste personal-social por el grupo de aplicación piloto, nos muestra la ausencia de diferencias significativas. Estos resultados vendrían a reforzar la contribución significativa de la aplicación del dispositivo EDPL en la consecución de las mejoras observadas en los procesos de ajuste relacionados con la lectura, de forma especial en el grupo de sujetos con mayores dificultades.

Con relación al valor explicativo diferencial de EDPL sobre el rendimiento

51. Las puntuaciones dinámicas obtenidas a raíz de la aplicación del dispositivo de ED, predicen de forma muy significativa el rendimiento en lectura, valorado por el profesorado. Así mismo, el valor predictivo del dispositivo EDPL con relación al rendimiento escolar, se ha mostrado muy superior a las diversas aproximaciones estáticas de evaluación de la lectura, así como de evaluación de la inteligencia general mediante una prueba de CI.

52. El dispositivo EDPL predice de forma significativa el rendimiento escolar, valorado por los tutores y por el profesorado de apoyo; así como el progreso de los sujetos, valorado igualmente por tutores y apoyo. Así mismo, la puntuación dinámica global, al igual que la puntuación dinámica de los procesos metacognitivos, también explican de forma significativa las puntuaciones de mejora. En este sentido, las puntuaciones dinámicas mostraron mayor valor de predicción que las puntuaciones pretest en ECO 1 con relación a las ganancias observadas, aportando de este modo información relevante sobre la validez diferencial de las puntuaciones dinámicas sobre las estáticas. Por su parte, la puntuación de mejora era la que mejor predecía el rendimiento escolar valorado por el profesorado de apoyo, aunque las puntuaciones dinámicas también contribuían de forma significativa a la explicación de estas valoraciones. Así mismo, las puntuaciones dinámicas EDPL fueron las que mejor predecían el rendimiento escolar valorado por los tutores para los subgrupos de *mejoradores* y *no mejoradores*.

53. La puntuación de mejora en la prueba de comprensión ECO 1 para el GE, predice de forma muy significativa el postest. Así mismo, predice significativamente el rendimiento escolar valorado por los profesores tutores y por los profesores de apoyo; así como el progreso de los sujetos, valorado por los profesores de apoyo. Por otra parte, la puntuación de mejora ha sido también la que mostró un mayor valor de predicción sobre el postest en esta prueba para el grupo de *mejoradores*, revelándose insuficiente la proporción de varianza explicada por el pretest para predecir el postest.

54. Las puntuaciones obtenidas en el pretest de la prueba de comprensión de primaria, no predicen la dirección de las diferencias observadas entre el pretest y el postest. La

divergencia entre los contrastes de la puntuación de mejora con las puntuaciones inicial y final de la prueba, en ausencia de diferencias significativas entre ambas aplicaciones; apuntan en esta dirección. De igual modo, la puntuación obtenida inicialmente en la prueba estática, no predice la actuación en el dispositivo de evaluación dinámica, ni así mismo, los cambios observados tras su aplicación con respecto a los resultados postratamiento. Estos datos vendrían a confirmar que la información otorgada por el pretest, resultaría a todas luces insuficiente con relación a la ofrecida por las dos puntuaciones establecidas como indicadores de la actuación en el dispositivo EDPL, es decir, la puntuación dinámica y la puntuación de mejora.

55. Las puntuaciones posttest y de mejora en la prueba de comprensión de secundaria (ECOS), explican la mayor parte del rendimiento escolar para el GC1; en cambio, son las puntuaciones dinámicas las que predicen de forma significativa el rendimiento de los sujetos del GE. Por otra parte, las puntuaciones estáticas (pretest y posttest) de la prueba ECOS no se han mostrado suficientes para explicar la mejora de los sujetos a raíz de la aplicación de EDPL, o la actuación de los sujetos durante las sesiones del dispositivo de evaluación dinámica.

56. La puntuación dinámica logra explicar y predecir de forma significativa el rendimiento escolar y el progreso experimentado por los sujetos del grupo de tratamiento en secundaria, valorados por los tutores y por el profesorado de apoyo. El pretest correlaciona negativamente y de forma significativa con el progreso de los sujetos, que presentó valores positivos. Estos datos avalan y confirman nuestra primera hipótesis experimental, con relación a la mejora del rendimiento en lectura de los sujetos con especiales dificultades de aprendizaje. Por otra parte, la puntuación dinámica obtuvo el nivel de correlación más alto con relación a las calificaciones finales en Lengua, aunque no lograba explicar de forma significativa estos valores; en cambio, no mostró valores notables de predicción con relación al posttest de la prueba criterio, ni tampoco con relación a la puntuación de mejora.

57. A juicio de los profesores tutores y de los aplicadores, los beneficios de la aplicación del dispositivo EDPL, habrían sido también para el grupo de sujetos que no logró ganancias postratamiento en la prueba criterio. Los aplicadores revelaron que un número importante de sujetos que no obtuvieron mejoras postratamiento, habría experimentado mejoras significativas en sus habilidades cognitivas, metacognitivas o de ajuste personal-social, aunque éstas mejoras no se vieran finalmente reflejadas proporcionalmente en las puntuaciones obtenidas en las pruebas criterio, o en las valoraciones efectuadas en las *Hojas de registro y evaluación*. Por otra parte, la mayoría de las valoraciones efectuadas por el profesorado tutor y de apoyo sobre el rendimiento escolar y sobre el progreso de los sujetos, así como sobre las calificaciones académicas, fueron positivas para los *no mejoradores*, situándose algunas de ellas por encima de las que obtuvieron los mejoradores. En este caso, el valor de las puntuaciones de mejora no se correspondió con las valoraciones del profesorado. Estas valoraciones habrían supuesto una actuación positiva de estos sujetos que no lograron mejoras, aunque esto no se vio reflejado finalmente en sus resultados en la prueba de comprensión. De este modo, las puntuaciones dinámicas han sido las que mejor predecían las valoraciones sobre el rendimiento escolar efectuadas por el profesorado para los no mejoradores.

58. Los análisis sobre el valor predictivo del dispositivo EDPL con relación a la escala de ajuste APSL, pusieron de manifiesto una relación muy significativa entre la puntuación dinámica global y la puntuación dinámica de los procesos de ajuste. Estos

resultados avalan la conexión entre el rendimiento general de los sujetos en el dispositivo EDPL y la puesta en práctica de habilidades relacionadas con los procesos de ajuste personal-social.

59. La puntuación dinámica global del dispositivo EDPL, así como la puntuación dinámica que hacía referencia exclusiva a los procesos de ajuste personal-social, mostraron un valor de predicción sobre el rendimiento escolar significativamente superior al que obtuvieron las puntuaciones pretest o postest, obtenidas en APSL. Las puntuaciones dinámicas lograban explicar significativamente el rendimiento y el progreso valorado por los tutores. Así mismo, la puntuación dinámica global predecía de forma significativa el rendimiento valorado por los aplicadores; y, junto con la puntuación dinámica de los procesos de ajuste, lograba explicar un porcentaje de varianza similar que el pretest, sobre el progreso experimentado por los sujetos.

60. La puntuación dinámica global y la de ajuste personal-social mostraron un valor de predicción superior al de las puntuaciones pretest o postest obtenidas en la escala APSL, con relación a la actitud de los sujetos ante la lectura; valorada ésta por tutores y profesores de apoyo. La puntuación dinámica global fue el único factor determinante en la explicación de las valoraciones efectuadas por parte de los tutores. Por su parte, la valoración de la actitud por parte de los aplicadores fue predicha de forma significativa por la puntuación dinámica de los procesos de ajuste; y la valoración del progreso reveló que era la puntuación pretest la variable que mejor predecía este valor, aunque el contraste del modelo no resultara significativo. También en esta ocasión, la conjunción de las puntuaciones dinámicas lograba explicar este valor de forma similar. Por otra parte, las puntuaciones postest del GE predicen por sí solas las calificaciones finales; al igual que también lo hace la conjunción de la puntuación dinámica global y la de ajuste; y la conjunción de la puntuación dinámica global y la puntuación de mejora, resultando ambos modelos igualmente significativos estadísticamente.

61. El nivel de partida en ajuste personal-social, medido a través de las puntuaciones iniciales de la escala APSL, conjuntamente con el CI pretest, no predice adecuadamente la puntuación postest en la escala, y se muestra insuficiente para explicar adecuadamente los resultados obtenidos tras la aplicación del dispositivo. Por su parte, el nivel de inteligencia en el pretest tampoco predice acertadamente el nivel de ajuste personal-social postratamiento. Sin embargo, cuando los análisis toman en consideración el segundo de los criterios formulados, el nivel de ajuste inicial predice de forma significativa el postest; aunque el porcentaje de varianza no explicada por el pretest, deja lugar a un posible efecto de la aplicación de EDPL, abriendo la posibilidad de obtener información adicional a la obtenida con la administración de la escala APSL.

62. La ausencia de tratamiento en el GC1 habría dado validez a la predicción inicial de la puntuación pretest en APSL sobre el rendimiento. El pretest figuraba como la única variable que predecía de forma significativa el rendimiento escolar y el progreso experimentado por los sujetos, valorado por los tutores. Las puntuaciones pretest del GC1 predecían el rendimiento escolar y las calificaciones escolares mejor que las obtenidas en el postest. De este modo, la valoración inicial en este grupo se mostraría suficiente para informar sobre la actuación de los sujetos. En el GE, la aplicación del dispositivo EDPL habría contribuido significativamente a la explicación del rendimiento escolar.

63. La puntuación de mejora en la prueba de CI predice el postest de forma muy significativa. El alto porcentaje de varianza explicada por la mejora (48%), unido a las ganancias postratamiento del GE, muestran la relevancia del proceso de evaluación dinámica y mediación llevado a cabo entre las fases pretest y postest del estudio, y pondría de manifiesto la contribución significativa del tratamiento aplicado. La actuación conjunta de las puntuaciones de mejora y las puntuaciones dinámicas (global y metacognitiva), lograba predecir también de forma significativa las puntuaciones postratamiento en Cattell. Estos resultados establecen diferencias notables con los que obtuvo el GC1, en el que las puntuaciones de mejora tenían un peso mucho menor en la explicación del postest. Por otra parte, la puntuación dinámica global y la puntuación dinámica de los procesos metacognitivos, incrementaron sus valores de correlación con la puntuación postratamiento CI, con relación a los valores obtenidos con el pretest. La puntuación dinámica global y la metacognitiva, también obtuvieron valores de correlación con la mejora, significativamente mayores que los obtenidos por la puntuación pretest. De igual forma, el CI incrementó en el postest su nivel de correlación con el rendimiento escolar, obteniendo valores significativos para la valoración del progreso efectuada por los tutores, y para la valoración del rendimiento efectuada por los aplicadores.

64. Las puntuaciones dinámicas (global y metacognitiva) han mostrado un valor de predicción sobre el rendimiento escolar, muy superior al de las puntuaciones CI. Las puntuaciones dinámicas predicen de forma muy significativa el rendimiento escolar, valorado por tutores y aplicadores; así como el progreso experimentado por los sujetos, valorado éste por los tutores. Así mismo, las puntuaciones dinámicas obtuvieron valores de correlación más altos con las calificaciones finales en el área de Lengua, que las puntuaciones CI. La puntuación dinámica de los procesos metacognitivos lograba también explicar de forma muy significativa el rendimiento escolar. Por su parte, la puntuación de mejora era la variable que mejor predecía el progreso experimentado por los sujetos, valorado por los aplicadores. En definitiva, las puntuaciones dinámicas, incluida la puntuación de mejora, fueron las que mostraron, para el GE, un valor de predicción mayor con relación al rendimiento escolar y al progreso de los sujetos, valorado por los profesores. En cambio, los valores de rendimiento, progreso y calificaciones del GC1, no eran explicados de forma significativa por ningún factor.

65. El valor de predicción del dispositivo EDPL con relación al rendimiento escolar, es ligeramente mayor para el grupo de sujetos con menores puntuaciones en la prueba de CI, lo que parece estar en consonancia con los resultados obtenidos por Guthke y Wingefeld (1992). En cualquier caso, este dato debe adoptar un carácter exploratorio, debido al escaso número de efectivos con puntuaciones medias y altas en CI. Por otra parte, hemos podido comprobar que la puntuación dinámica lograba explicar las valoraciones sobre el rendimiento en los sujetos con bajas puntuaciones en comprensión y puntuaciones medias en CI, aunque los valores de correlación fueran menores. De esta forma, no se descarta que el valor predictivo de las puntuaciones dinámicas EDPL, pueda aplicarse a los sujetos con puntuaciones medias o altas en CI.

66. La puntuación dinámica EDPL, extraída del proceso de aplicación del dispositivo, ha mostrado valores de predicción sobre el rendimiento en lectura, el ajuste personal-social, así como sobre el progreso de los sujetos; significativamente mayores que los que mostraron las puntuaciones obtenidas de la aplicación de pruebas estáticas, ya fueran éstas pruebas relacionadas con el contenido a evaluar (las pruebas de comprensión y las de ajuste personal-social), o pruebas generales de aptitud cognitiva.

Referencias Bibliográficas ¹

- Adams, M. J. (1980): Failures to comprehend and levels of processing in reading. En R. J. Spiro; B. C. Bruce y W. F. Brewer (eds.), *Theoretical issues in reading comprehension*. Hillsdale, N.J.: Lawrence Erlbaum.
- Adams, M.J. (1994): Modelling the connections between word recognition and reading. En R.B. Ruddell; M.R. Ruddell y H. Singer (eds.), *Theoretical models and processes of reading* (4ª ed., pp. 838-863). Newark, DE: International Reading Association.
- Adams, M.J. y Collins, A. (1977): *A schema-theoretic view of reading* (Tech. Report nº 32). Urbana: University of Illinois, Center for the Study of Reading.
- Aguilera, A. (1997): *Evaluación de las habilidades de pensamiento en situaciones de interacción social*. Tesis doctoral. Universidad de Sevilla.
- Aguilera, A. y García, I. (2004): Evaluar interviniendo, intervenir evaluando: una propuesta de actuación ante las dificultades del aprendizaje. *Apuntes de Psicología*, vol. 22 (3), págs. 309-322.
- Aguilera, A. y Mora, J. (2003): Pautas de interacción en el aula y desarrollo cognitivo: los Patrones de Movilización Cognitiva (PMC). *Apuntes de Psicología*, vol. 21(1), págs. 5-27.
- Aguilera, A. y Rodríguez, I. (2003): La intervención ante dificultades del aprendizaje. En A. Aguilera (coord.), *Introducción a las dificultades del aprendizaje*. Madrid: McGraw-Hill.
- Aguilera, A. y Saldaña, D. (2003): La evaluación de las dificultades del aprendizaje. En A. Aguilera (coord.), *Introducción a las dificultades del aprendizaje*. Madrid: McGraw-Hill.
- Alegría, J. (1985): Por un enfoque psicolingüístico del aprendizaje de la lectura y sus dificultades. *Infancia y Aprendizaje*, nº 29, págs. 79-94.
- Alegría, J. (2006): Por un enfoque psicolingüístico del aprendizaje de la lectura y sus dificultades —20 años después—. *Infancia y Aprendizaje*, nº 29 (1), págs. 93-111.
- Alexander, J.M.; Carr, M. y Schwanenflugel, P.J. (1995): Development of metacognition in gifted children: Directions for future research. *Developmental Review*, 15, págs. 1-37.
- Alliende, F.; Condemarín, M. y Milicic, N. (1991): *Prueba CLP. Prueba de comprensión lectora de complejidad lingüística progresiva*. Madrid: CEPE.
- Alonso Tapia, J. (1992): Evaluación del potencial de cambio intelectual, aptitudinal y de aprendizaje. En: R. Fernández-Ballesteros, *Introducción a la evaluación psicológica* (1), págs. 454-495. Madrid: Pirámide.
- Alonso Tapia, J. (1995): La evaluación de la comprensión lectora. *Textos de Didáctica de la Lengua y de la Literatura*, 5, 63-78.
- Alonso Tapia, J. Y Cols. (1992): *Leer, Comprender y Pensar: Nuevas estrategias y técnicas de evaluación*. Madrid: CIDE.
- Alonso Tapia, J.; Carriedo, N. y González, E. (1992): Evaluación de la capacidad de comprender y resumir lo importante: la batería IDEPA. En J. Alonso Tapia y Cols., *Leer, Comprender y Pensar. Nuevas estrategias y técnicas de evaluación* (págs. 59-183). Madrid: Servicio de Publicaciones del Ministerio de Educación y Ciencia. C.I.D.E.
- Alonso Tapia, J.; Carriedo, N. y Mateos, M.M. (1992): Evaluación de la supervisión y regulación de la comprensión. La batería SURCO. En J. Alonso Tapia y Cols., *Leer, Comprender y Pensar. Nuevas estrategias y técnicas de evaluación* (págs. 11-57), Madrid: Servicio de Publicaciones del Ministerio de Educación y Ciencia. C.I.D.E.
- Alonso Tapia, J. y Carriedo, N. (1996): Problemas de comprensión lectora: Evaluación e intervención. En C. Monereo e I. Solè (coords.), *El asesoramiento psicopedagógico: una perspectiva profesional y constructivista*. Madrid. Alianza.
- Alonso Tapia, J. y Gutiérrez, F. (1992): De la comprensión al pensamiento crítico: la batería CRITEX. En J. Alonso Tapia y Cols., *Leer, Comprender y Pensar. Nuevas estrategias y técnicas de evaluación* (págs. 185-287). Madrid: Servicio de Publicaciones del Ministerio de Educación y Ciencia. C.I.D.E.
- Alonso Tapia, J. y Mateos Sanz, M. (1985): Comprensión lectora: Modelos, entrenamiento y evaluación. *Infancia y Aprendizaje*, 31-32, págs. 5-19.

¹ Hemos incluido aquí también las referencias que se recogen en el volumen de anexos

- Alonso Tapia, J. y Montero, I. (2001): Orientación motivacional y estrategias motivadoras en el aprendizaje escolar. En C. Coll; J. Palacios y A. Marchesi (comps.), *Desarrollo Psicológico y Educación. 2. Psicología de la educación escolar* (págs. 259-283). Madrid: Alianza.
- Anderson, R.C. y Pichert, J.W. (1977): *Recall of previously unrecallable information following a shift in perspective* (Inf. Tec. n° 41). Urbana, Universidad de Illinois, Center for the Study of Reading, abril 1977 (ed. 142 974).
- Aranov, Z. (1999): *Validity and Reliability of the ACFS Behavior Observation Scale*. Unpublished masters thesis; Touro College, New York, NY. ERIC document TM 030602.
- Armbruster, B.B.; Anderson, T.H. y Ostertag, J. (1987): Does text structure/summarization instruction facilitate learning from expository text? *Reading Research Quarterly*, XXIV, 331-346.
- Ashman, A. y Conway, R.N.F. (1989): *Cognitive strategies for special education*. Londres: Routledge.
- August, D.L.; Flavell, J.H. y Clift, R. (1984): Comparaison of comprehension monitoring of skilled and less skilled readers. *Reading Research Quarterly*, XX, 39-53. Traducción en *Infancia y Aprendizaje*, 31-32, Una comparación del control de comprensión en lectores más y menos competentes, 1985, págs. 129-143.
- Ausubel, D.P. (1963): *The psychology of meaningful verbal learning*. New York: Grune y Stratton.
- Ausubel, D.P.; Novak, J.D. y Hanesian, H. (1978): *Educational psychology: a cognitive view*. New Cork: Holt. (Trad. cast.: *Psicología educativa: un punto de vista cognoscitivo*. México: Trillas, 1983).
- Aznar, E.; Cross, A. y Quintana, L. (1993): Lectura y coherencia textual (análisis de un texto narrativo). *Comunicación, Lenguaje y Educación*, 17, 15-28.
- Baddeley, A. (1982): *Your memory*. Londres: Sidgewick and Jackson. (Trad. cast. de M.V. Sebastián y T. del Amo: *Su memoria. Cómo conocerla y dominarla*. Madrid: Debate, 1984).
- Baddeley, A.D. (1996): Exploring the central executive. *Quarterly Journal of Experimental Psychology*, 49a, 5-28.
- Baker, J. y Feuerstein, R. (1991): The Modifying Environment and Other Environmental Perspectives in Group Care: A Conceptual Contrast and Integration. *Residential Treatment for Children and Youth*, Vol. 8(3). New York: Haworth Press.
- Baker, L. (1989): Metacognition, comprehension monitoring, and the adult reader. *Educational Psychology Review*, 1, 3-38.
- Baker, L. y Brown, A. (1984): Cognitive monitoring in reading. En J. Flood (ed.), *Understanding Reading Comprehension: Cognition and human performance*. Delaware: International Reading Association.
- Baumann, J.F. (1985): La eficacia de un modelo de instrucción directa en la enseñanza de la comprensión de ideas principales. *Infancia y Aprendizaje*, 31-32, 89-107.
- Baumann, J.F. (1986): *Teaching main idea comprehension*. Delaware: IRA.
- Beckmann, J.F. y Guthke, J. (1999): *Psychodiagnostik des schlussfolgernden Denkens [Assessing reasoning ability]*. Göttingen: Hogrefe.
- Beretvas, T. (1999): *A conceptual framework for the validation process*. Unpublished manuscript. University of Washington.
- Bethge, H.J.; Carlson, J. y Wiedl, K.H. (1982): The effects of dynamic assessment procedures on Raven Matrices performance, visual search behaviour, test anxiety, and test orientation. *Intelligence*, 6, 89-97.
- Bickman, L. y Rog, D.J. (1998): Introduction: Why a handbook of applied social research methods? En: L. Bickman y D.J. Rog (ed.) *Handbook of applied social research methods*. Thousand Oaks (CA), Sage, págs. IX-XIX.
- Blanton, W.; Wood, K. y Moorman, G. (1990): The role of purpose in reading instruction. *The Reading Teacher*, 43 (7), 486-493.

- Bleichrodt, N.; Drenth, P.J.D.; Zaal, J.N. y Resing, W.C.M. (1984): *Revisie Amsterdamse Kinder Intelligentie Test* [Revision Amsterdam Child Intelligence Test]. Lisse: Swets & Zeitlinger.
- Bofarull, M.T. (2001): Evaluación de la comprensión lectora. Propuesta de una pauta de observación. En: VVAA, *Comprensión lectora. El uso de la lengua como procedimiento*. Barcelona: Graò.
- Borkowski, J.G.; Weyhing, R.S. y Turner, L.A. (1986): Attributional retraining and the teaching of strategies. *Exceptional Children*, 53, 130-137.
- Botella, J. y Barriopedro, M.I. (1995): Análisis de datos. En R. Fernández-Ballesteros (ed.), *Evaluación de programas. Una guía práctica en ámbitos sociales, educativos y de salud*, págs. 173-203.
- Botella, J.; León, O.; San Martín, R. y Barriopedro, M.I. (2001): *Análisis de Datos en Psicología I*. Madrid: Pirámide.
- Bradley, T.B. (1983): Remediation of cognitive deficits: A critical appraisal of the Feuerstein model. *Journal of Mental Deficiency Research*, 27, págs. 79-92.
- Bransford, J.D.; Vye, N.J.; Adams, L.T. y Perfetto, G.A. (1989): Learning skills and acquisition of knowledge. En A. Lesgold y R. Glaser (eds.), *Foundation for a Psychology of education* (págs. 199-249). Hillsdale, N.J.: Lawrence Earlbaum Associates.
- Bronfenbrenner, U. (1987): *La ecología del desarrollo humano*. Barcelona: Paidós.
- Brooks, L.W. y Danserau, D.F. (1983): Effects of structural schema training and Text Organization on Expository Prose Processing. *Journal of Educational Psychology*, 75, 6, 811-820.
- Brooks, N.D. (1997): *An exploratory study into the cognitive modifiability of pre-school children using dynamic assessment*. Unpublished masters thesis; University of Newcastle-Upon-Tyne; Newcastle, United Kingdom.
- Brown, A.L. (1980): Metacognitive development en reading. En R. Spiro, B. Bruce y W. Brewer (eds.): *Theoretical issues in reading comprehension*. Hillsdale, New Jersey: L. Erlbaum.
- Brown, A.L. (1985): Metacognition: the development of selective attention strategies for learning from texts. En H. Singer y R. Ruddell (eds.), *Theoretical Models and Processes of Reading* (págs. 501-526). Delaware, N.J.: I.R.A.
- Brown, A.L. (1987): Metacognition, executive control, self-regulation, and other more mysterious mechanisms. En: E. Weinert y R. Kluwe (eds.). *Metacognition, motivation, and understanding*, págs. 65-116. Hillsdale, N.J.: Erlbaum.
- Brown, A.L.; Bransford, J.D.; Ferrara, R.A. y Campione, J.C. (1983): Learning, remembering and understanding. En P.H. Mussen (ed.), *Handbook of child psychology, vol. 3*: J. H. Flavell y E. Markman (eds.), *Cognitive development*, págs. 77-166. Nueva York: Wiley.
- Brown, A.L. y Campione, J.C. (1981): Inducing flexible thinking: A problem of access. En: M. Friedman; J.P. Das y N. O'Connor (comps.), *Intelligence and learning*. New York: Plenum Press.
- Brown, A.L. y Campione, J.C. (1984): Three faces of transfer: Implications for early competence, individual differences, and instruction. En M.E Lamb; A.L. Brown y B. Rogoff (eds.), *Advances in developmental psychology. Vol. 3*, págs. 143-192. Hillsdale, N.J.: Lawrence Erlbaum.
- Brown, A.L.; Campione, J.C.; Bray, N.W. y Wilcox, B.L. (1973): Keeping track of changing variables: Effects of rehearsal training and rehearsal prevention in normal and retarded adolescents. *Journal of Experimental Psychology*, 101, 123-131.
- Brown, A.L.; Day, J.D. y Jones, R.S. (1983): The development of plans for summarizing texts. *Child Development*, 54, 968-979.
- Brown, A.L. y Ferrara, R.A. (1985): Diagnosing zones of proximal development. En J.V. Wertsch (comp.), *Culture, communication, and cognition: Vygotskian perspectives*. New York: Cambridge University Press, págs. 273-305
- Brown, A.L. y Palinscar, A.S. (1989): Guided, cooperative learning and individual knowledge acquisition. En L.B. Resnick (ed.), *Knowledge, learning, and instruction: Essays in honor of Robert Glaser*, págs. 393-451. Hillsdale, N.J.: Erlbaum.

- Bruininks, R.H.; Hill, B.K.; Weatherman, R.F.; y Woodcock, R.W. (1986): *ICAP. Inventory for client and agency planning. Examiner's manual*. Allen, DLM Teaching Resources.
- Bruner, J.S. (1990): *Acts of meaning*. Cambridge, MA: Harvard University Press.
- Bruning, R.H.; Schraw, G.J.; y Ronning, R.R. (2002): *Psicología cognitiva e instrucción*. Madrid: Alianza Editorial.
- Bücheel, F.P. y Scharnhorst, U. (1993): The Learning Potential Testing Device (LPAD): Discussion of theoretical and methodological problems. En J.H.M. Hamers; K. Sijtsma y A.J.J.M. Ruijsenaars (comps.), *Learning potential testing*. Amsterdam: Swets and Zeitlinger, págs. 83-111.
- Budoff, M. (1967): Learning potential among institutionalized young adult retardates. *American Journal of Mental Deficiency*, 72, págs. 404-411.
- Budoff, M. (1968): Learning potential as a supplementary testing procedure. En J. Hellmuth (comp.), *Learning disorders*. Seattle: WA, Special Child, vol. 3, págs. 295-343.
- Budoff, M. (1975): *Learning potential among educable retarded pupils*, informe definitivo. Research Institute for Educational Problems, 29, Ware Street. Cambridge, MA.
- Budoff, M. (1987a): The validity of learning potential assessment. En C.S. Lidz (comp.), *Dynamic assessment: An interactional approach to evaluating learning potential*. New York: Guilford Press, págs. 52-81.
- Budoff, M. (1987b): Measures for assessing learning potential. En: C.S. Lidz (comp.). *Dynamic assessment: An interactional approach to evaluating learning potential*. New York: Guilford Press, págs. 173-195.
- Budoff, M.; Corman, L. y Gimon, A. (1976): An educational test of learning potential testing with Spanish-speaking youth. *Interamerican Journal of Psychology*, 10, págs. 13-24.
- Budoff, M. y Friedman, M. (1964): «Learning potential» as a testing approach to the adolescent mentally retarded. *Journal of Consulting Psychology*, 28, págs. 434-439.
- Budoff, M. y Hamilton, J. (1976): Optimizing test performance of the moderately and severely mentally retarded. *American Journal of Mental Deficiency*, 81, págs. 49-57.
- Budoff, M. y Hutten, L. (1971): The development of learning potential measures based on Raven's. *Studies in Learning Potential*, 14. Cambridge: RIEP.
- Budoff, M.; Meskin, J. y Harrison, R.G. (1971): An educational test of the learning potential hypothesis. *American Journal of Mental Deficiency*, 76, págs. 159-169.
- Buron, J. (1993): *Enseñar a aprender: Introducción a la metacognición*. Bilbao: Mensajero.
- Campbell, D. y Stanley, J. (1966): *Experimental y Quasi-experimental designs for research*. Chicago: Rand McNally Collage Publishing Company.
- Campbell, D. y Stanley, J. (1978): *Diseños experimentales y cuasiexperimentales en la investigación social*. Buenos Aires: Amorrortu Editores.
- Campione, J.C. (1987): Metacognitive components of instructional research with problem learners. En F.E. Weinert y R.H. Kluwe (eds.), *Metacognition, motivation and understanding*. Hillsdale: Erlbaum.
- Campione, J.C. (1989): Assisted testing: A taxonomy of approaches and an outline of strengths and weaknesses. *Journal of Learning Disabilities*, 22 (3), págs. 151-165.
- Campione, J.C. y Brown, A.L. (1977): Memory and metamemory development in educable retarded children. En R.V. Kail y J.W. Hagen (eds.), *Perspectives on the development of memory and cognition*. Hillsdale, N.J.: Lawrence Earlbaum Associates.
- Campione, J.C. y Brown, A.L. (1979): *Human intelligence*. Norwood, N.J.: Albex.
- Campione, J.C. y Brown, A.L. (1987): Linking dynamic testing with school achievement. En C.S. Lidz (comp.), *Dynamic assessment: An international approach to evaluating learning potential*, págs. 82-115. New York: Guilford Press.
- Campione, J.C.; Brown, A.L. y Bryant, N.R. (1985): Individual differences in learning and memory. En R.J. Sternberg (comp.), *Human abilities: An information-processing approach*, págs. 103-126. New York: Freeman
- Campione, J.C.; Brown, A.L. y Ferrara, R.A. (1987): Retraso mental e inteligencia. En R. Sternberg: *Inteligencia humana, Vol. II*. Barcelona: Paidós. Original en inglés, 1982.

- Campione, J.C.; Brown, A.L.; Ferrara, R.A. y Bryant, N.R. (1984): The zone of proximal development: Implications for individual differences and learning. En B. Rogoff y J.V. Wertsch (eds.), *Children's learning in the zone of proximal development. New directions for child development*, 23, págs. 77-91. San Francisco: Jossey-Bass.
- Campione, J.C.; Brown, A.L.; Ferrara, R.A.; Jones, R.S. y Steinberg, E. (1985): Differences between retarded and not-retarded children in transfer following equivalent learning performance: Breakdowns in flexible use of information. *Intelligence*, 9, págs. 297-315.
- Campllonch, J.M. y Fernández-Ballesteros, R. (1981): Evaluación del potencial de aprendizaje. En R. Fernández-Ballesteros (ed.) *Nuevas aportaciones en evaluación conductual*, págs. 257-278. Valencia: Alphaplus.
- Carey, S. y Spelke, E. (1994): Domain specific knowledge and conceptual change. En L. Hirschfeld y S. Gelman (eds.), *Mapping the mind*. Cambridge, Massachusetts: Cambridge University Press.
- Carlson, J.S. y Das, J.P. (1997): A process-approach to remediating word-decoding deficiencies in Chapter I Children. *Learning Disabilities Quarterly*.
- Carlson, J.S. y Wiedl, K.H. (1976): The factorial análisis of perceptual and abstract reasoning abilities in tests of concrete operational thought. *Educational and Psychological Measurement*, 36, págs. 1015-1019.
- Carlson, J.S. y Wiedl, K.H. (1978): Use of testing-the-limits procedures in the testing of intellectual capabilities in children with learning difficulties, *American Journal of Mental Deficiency*, 11, págs. 559-564.
- Carlson, J.S. y Wiedl, K.H. (1979): Toward a differential testing approach: Testing-the-limits employing the Raven matrices. *Intelligence*, 3, págs. 323-344.
- Carlson, J.S. y Wiedl, K.H. (1980): Applications of a dynamic testing approach in intelligence assessment: Empirical results and empirical formulations. *Zeitschrift Fuer Differentielle und Diagnostische Psychologie*, 1, 303-318.
- Carlson, J.S. y Wiedl, K.H. (1992a): Principles of dynamic testing: The application of a specific model. *Learning and Individual Differences*, 4, págs. 153-166.
- Carlson, J.S. y Wiedl, K.H. (1992b): The dynamic testing of intelligence. En H.C. Haywood y D. Tzuriel (comps.), *Interactive testing*, New York: Springer-Verlag, págs. 167-186.
- Carlson, J.S. y Wiedl, K.H. (2000): The validity of dynamic assessment. En C.S. Lidz y J.G. Elliot (comps.), *Dynamic Assessment: Prevailing Models and Applications* (6), págs. 681-712. Greenwich, CT, Elsevier-JAI.
- Carriedo, N. (1996): Consideraciones metodológicas sobre los programas de instrucción en comprensión de las ideas principales y sobre su eficacia. *Infancia y Aprendizaje*, 73, págs. 87-107.
- Carriedo, N. y Alonso Tapia, J. (1994): *¿Cómo enseñar a comprender un texto?* Madrid: Cuadernos del ICE, 10. Ediciones de la Universidad Autónoma de Madrid.
- Carrillo, M.S. (1993): *Desarrollo de la conciencia fonológico-silábica y adquisición de la lectura*. Tesis Doctoral. Universidad de Murcia.
- Carroll, J.B. y Sapon, S.M. (1958): *Modern Language Aptitude Test*. New York: Psychological Corporation.
- Carver, R. (1974): Two dimensions of tests: Psychometric and edumetric. *American Psychologist*, 29, págs. 512-518.
- Castro García, R. (1993): *Estrategias de Metacompreensión en alumnos de 2º y 3º de E.G.B. y su influencia en el rendimiento en Ciencias Sociales*. Tesis Doctoral. Madrid: UNED.
- Català, G.; Català, M.; Molina, E. y Monclús, R. (2001): *Evaluación de la comprensión lectora. Pruebas ACL (1º-6º de primaria)*. Barcelona: Graó.
- Cattell, R.B. y Cattell, A.K.S. (1974/2001): *Test de factor «G» (Escala 2 y 3)*. Madrid: TEA.
- Cavanaugh, J.C. y Perlmutter, M. (1982): Metamemory: a critical examination. *Child Development*, 53, 11-28.
- Chi, M.T.H. (1981): Interactive roles of Knowledge and strategies in development. En J. S. Chipman y R. Glaser (eds.), *Thinking and learning skills*. Hillsdale, N.J.: Lawrence Earlbaum Associates.

- Chi, M.T.H. (1987): Representing Knowledge and metaknowledge: implications for interpreting metamemory research. En F.E. Weinert y R.H. Kluwe (eds.), *Metacognition, motivation and understanding*. Hillsdale: Erlbaum.
- Cioffi, G. y Carney, J. (1983): Dynamic assessment of reading disabilities. *Reading Teacher*, 36, 764-768.
- Cole, M. (1985): The zone of proximal development: where culture and cognition create each other. En J.V. Wertsch (comp.), *Culture, communication, and cognition: Vygotskian perspectives*, págs. 146-161, New York: Cambridge University Press.
- Cole, M. (1990): Cognitive development and formal schooling. En L. Moll (comp.), *Vygotski and education. Instructional implications and applications of sociohistorical psychology*, págs. 89-110. New York: Cambridge University Press.
- Cole, M. y D'Andrade, R. (1982): The influence of schooling on concept formation: some preliminary conclusions. En *The Quarterly Newsletter of LCHC*, 4, 19-26.
- Cole, M.; Gay, J.; Glick, J. y Sharp, D. (1971): *The cultural contexts of learning and thinking*. New York: Basic Books.
- Cole, M. y Scribner, S. (1974): *Culture and thought: a psychological interpretation*, New York: John Wiley. (Trad. cast. *Cultura y pensamiento*, Limusa, 1977)
- Cole, M. y Scribner, S. (1996): Introducción de *El desarrollo de los procesos psicológicos superiores*, de L.S. Vygotski. Barcelona: Grijalbo Mondadori.
- Coll, C. (2001): Constructivismo y educación: la concepción constructivista de la enseñanza y el aprendizaje. En C. Coll; J. Palacios y A. Marchesi (comps.), *Desarrollo Psicológico y Educación. 2. Psicología de la educación escolar* (págs. 157-186). Madrid: Alianza.
- Coll, C.; Barberà, E. y Onrubia, J. (2000): La atención a la diversidad en las prácticas de evaluación. *Infancia y Aprendizaje*, 90, 111-132.
- Coll, C. y Edwards, D. (eds.) (1996): *Enseñanza, aprendizaje y discurso en el aula. Aproximaciones al discurso educacional*. Madrid: Fundación Infancia y Aprendizaje.
- Coll, C.; Martín, E. y Onrubia, J. (2001): La evaluación del aprendizaje escolar: dimensiones psicológicas, pedagógicas y sociales. En C. Coll; J. Palacios y A. Marchesi (comps.), *Desarrollo Psicológico y Educación. 2. Psicología de la educación escolar* (págs. 549-567). Madrid: Alianza.
- Coll, C. y Miras, M. (2001): Diferencias individuales y atención a la diversidad en el aprendizaje escolar. En C. Coll; J. Palacios y A. Marchesi (comps.), *Desarrollo Psicológico y Educación. 2. Psicología de la educación escolar* (págs. 331-351). Madrid: Alianza.
- Coll, C. y Onrubia, J. (1999): Evaluación de los aprendizajes y atención a la diversidad. En C. Coll (coord.), *Psicología de la instrucción: la enseñanza y el aprendizaje en la educación secundaria*, págs. 141-168. Barcelona: Horsori/ICE UB.
- Coll, C. y Rivièrè, A. (1985): Individuación e interacción en el periodo sensoriomotor: apuntes sobre la construcción genética del sujeto y el objeto social. Xxème Journées d'Etude de l'APSLF. Lisboa.
- Coll, C. y Solè, I. (1993): Los profesores y la concepción constructivista. En C. Coll; E. Martí; T. Mauri y Otros, *El constructivismo en el aula*, Barcelona: Graó, págs. 7-23.
- Coll, C. y Solè, I. (2001): Enseñar y aprender en el contexto del aula. En C. Coll; J. Palacios y A. Marchesi (comps.), *Desarrollo Psicológico y Educación. 2. Psicología de la educación escolar* (págs. 357-383). Madrid: Alianza.
- Colomer, T. y Camps, A. (1991): *Enseñar a llegir, ensenyar a comprendre*. Barcelona: Rosa Sensat/Edicions 62. (Trad. Cast. *Enseñar a leer, enseñar a comprender*. Madrid: Celeste/MEC, 1996)
- Colomina, R.; Onrubia, J. y Rochera, M.J. (2001): Interactividad, mecanismos de influencia educativa y construcción del conocimiento en el aula. En C. Coll; J. Palacios y A. Marchesi (comps.), *Desarrollo Psicológico y Educación. 2. Psicología de la educación escolar* (págs. 437-456). Madrid: Alianza.
- Cook, L.K. y Mayer, R.E. (1988): Teaching readers about structure of scientific text. *Journal of Educational Psychology*, 80, 488-546.

- Cook, T.D. y Reichardt, C.S. (1986): *Métodos cualitativos y cuantitativos en investigación evaluativa*. Madrid: Editorial Morata.
- Cook, T.S. y Campbell, D.T. (1976): The design and conduct of quasi experiments and true experiments in field settings. En M.D. Dunnette y L.M. Hough (eds.), *Handbook of Industrial and Organizational Psychology, vol. I*. New York: Raul MacNally.
- Cook, T.S. y Campbell, D.T. (1979): *Quasiexperimentation. Design and análisis issues for field settings*. Boston: Houghton Mifflin Company.
- Cooper, J.D. (1990): *Cómo mejorar la comprensión lectora*. Madrid. Aprendizaje Visor.
- Cormier, P.; Carlson, J.S. y Das, J.P. (1990): Planning ability and cognitive performance: The compensatory effects of a dynamic testing approach. *Learning and Individual Differences, 2*, págs. 437-449.
- Cubero, R. y Luque, A. (2001): Desarrollo, educación y educación escolar: la teoría sociocultural del desarrollo y del aprendizaje. En C. Coll; J. Palacios y A. Marchesi (comps.), *Desarrollo Psicológico y Educación. 2. Psicología de la educación escolar* (págs. 137-153). Madrid: Alianza.
- Cuetos, F. (1990): *Psicología de la lectura*. Madrid: Escuela Española.
- Cuetos, F.; Rodríguez, B. y Ruano, E. (1996): *PROLEC: Evaluación de los procesos lectores*. Madrid: TEA ediciones.
- Dansereau, D.F. (1979): Development and evaluation of learning strategy training program. *Journal of Educational Psychology, 71*, 64-73.
- Das, J.P. (1985): Remedial training for the amelioration of cognitive deficits in children. En A.F. Ashman y R.S. Laura (eds.), *The education and training of the mentally retarded: recent advances*. Londres: Croom Helm.
- Das, J.P.; Naglieri, J.A. y Kirby, J.R. (1994): *Assessment of cognitive processes: The PASS theory of intelligence*. Boston: Allyn and Bacon.
- Day, J.D. y Zajakowski, A. (1991): Comparisons of learning ease and transfer propensity in poor and average readers. *Journal of Learning Disabilities, 24*, págs. 421-428.
- De Coster, W. (1977): *Educación compensatoria*. Madrid: Servicio de publicaciones del M.E.C.
- De la Cruz, M.V. (1988): *Pruebas de lectura, niveles 1 y 2 (LECTURA)*. Madrid: TEA.
- De la Cruz, M.V. (1995): *Evaluación de la comprensión lectora*. Madrid: TEA.
- De la Cruz, M.V. (2000): *Prueba de Aptitudes en Educación Infantil (AEI)*. Madrid: TEA.
- De Vega, M. (1984): *Introducción a la psicología cognitiva*. Madrid: Alianza.
- Defior Citoler, S. y Ortúzar Sanz, R. (1993): La lectura: evaluación e intervención educativa. En R. Bautista (comp.), *Necesidades Educativas Especiales*. Archidona-Málaga: Aljibe.
- Del Río, P. (1999): El cambio histórico-cultural y las identidades de género: los jóvenes españoles ante la diferencia de sexo. *Cultura y Educación, 14-15*, 147-176.
- Delclos, V.R.; Vye, N.J.; Burns, M.S.; Bransford, J.D. y Hasselbring, T.S. (1992): Improving the quality of instruction: Roles for dynamic testing. En H.C. Haywood y D. Tzuriel (comps.), *Interactive testing*. New York: Springer-Verlag, págs. 317-332.
- Díez Fernández, E. (1980): *Test de comprensión verbal (TCV)*. Madrid: Publicaciones ICCE.
- Dillon, R.F. y Carlson, J.S. (1978): Testing for competence in three ethnic groups. *Educational and Psychological Measurement, 38*, págs. 436-443.
- Dixon, W.J. y Massey, F.J. (Jr) (1969): *Introducción al análisis estadístico*. Madrid: Ediciones del Castillo.
- Dockrell, J. y Mcshane, J. (1992): *Children's learning difficulties. A cognitive approach*. Cambridge, Massachusetts: Blackwell Publishers. En Cast. (1997) *Dificultades de aprendizaje en la infancia. Un enfoque cognitivo*. Barcelona: Paidós.
- Doerr, P. y Jensen, M.R. (1998): *Conceptual and applied aspects of the Teacher as Mediator program*. Paper presented at the meeting of the California Association for Mediated Learning. Oxnard, CA.
- Dole, J.; Brown, K. y Trathen, W. (1996): The effects of strategy instruction on the comprehension performance of at-risk students. *Reading Research Quarterly, 31* (1), págs. 62-68.
- Domínguez, A. y Cuetos, F. (1992): Desarrollo de las habilidades de reconocimiento de palabras en niños con distintas competencias lectoras. *Cognitiva, 4* (2), 193-208.

- Donoso, T. (1994): Evaluación de la lectura mediante pruebas estandarizadas. Análisis crítico y valoración de pruebas de lectura. En F. Cabrera; T. Donoso y M^a A. Marín: *El proceso lector y su evaluación*. Barcelona: Laertes.
- Dubois, M.E. (2003): *Una historia personal*. Ponencia presentada en las IX Jornadas Nacionales «Historia, usos y aprendizaje del lenguaje escrito». Sevilla, 2003.
- Dunn, L. y Dunn, L. (1981): *Peabody Picture Vocabulary Test-Revised*. Circle Pines, MN, American Guidance Service (Trad. cast. *TVIP: Test de vocabulario en imágenes: Peabody*. Madrid: Material y Ediciones Psicotécnicas, 1986).
- Edwards, D. y Mercer, N. (1988): *El conocimiento compartido. El desarrollo de la comprensión en el aula*. Barcelona Paidós/MEC.
- Elboj Saso, C.; Puigdelívol Aguade, I.; Soler Gallart, M. y Valls Carol, R.(2002): *Comunidades de aprendizaje. Transformar la educación*. Barcelona: Graó.
- Elliot, S.N. y Fuchs, L.S. (1997): The utility of curriculum-based measurement and performance assessment as alternatives to traditional intelligence and achievement tests. *School Psychology Review*, 26, 224-233.
- Ellis, A. (1984): *Reading, writing and dyslexia. A cognitive analysis*, Londres: LEA.
- Embretson, S.E. (1987): Toward development of a psychometric approach. En C. Lidz (ed.). *Dynamic assessment*. New York: Guilford Press, págs. 141-172.
- Embretson, S.E. y Reise, S.R. (2000): *Item response theory for psychologists*. Mahwah, N.J.: Erlbaum Publishers.
- Emerson, L. (1985): *To sing our own songs: Cognition and culture in Indian education*. Report from a workshop for American Indian educators on the Learning Potential Assessment Device and Instrumental Enrichment programs. Shiprock, Navajo Nation, N.M. New York: Association of American Indian Affairs in cooperation with the NiHa'Alchini Ba Educational Center, Shiprock, N.M.
- Emerson, L. (1986): *August. Feuerstein's MLE and American Indian Education*. Paper presented at Mediated Learning Experience International Workshop. Jerusalem, Israel.
- Ericson, K.A. y Simon, H.A. (1980): Verbal reports as data. *Psychological Review*, 87, 215-251.
- Fernández-Ballesteros, R. (1980): *Psicodiagnóstico. Concepto y metodología*. Madrid: Cincel-Kapelusz.
- Fernández-Ballesteros, R. (1992): La observación. En R. Fernández-Ballesteros, *Introducción a la evaluación psicológica I*, págs. 137-182. Madrid: Pirámide
- Fernández-Ballesteros, R. (1996): Cuestiones de fiabilidad y validez en la evaluación del potencial de aprendizaje. En S. Molina y M. Fandos (comps.). *Educación cognitiva, vol. I*. Zaragoza: Mira Pub.
- Fernández-Ballesteros, R. y Calero, M.D. (1993): Measuring learning potential. *International Journal of Cognitive Education and Mediated Learning*, 3, 9-20.
- Fernández-Ballesteros, R. y Calero, M.D. (1995): Training effect on intelligence of older persons. *Archives of Gerontology and Geriatrics*, 20, 135-148.
- Fernández-Ballesteros, R. y Calero, M.D. (2000): The assessment of learning potential: the EPA instrument. En C.S. Lidz y J.G. Elliott (comps.), *Dynamic Assessment: Prevailing Models and Applications, vol. 6*, págs. 293-323, Greenwich, CT, Elsevier-JAI.
- Fernández-Ballesteros, R.; Calero, M.D.; Campllonch, J.M. y Belchi, J. (1987/1990): *EPA: Evaluación del potencial de aprendizaje*. Madrid: MEPSA.
- Fernández-Ballesteros, R.; Calero, M.D.; Campllonch, J.M. y Belchi, J. (2000): *Evaluación del potencial de aprendizaje (EPA)*. Madrid: TEA ediciones.
- Fernández-Ballesteros, R.; Campllonch, J.M. y Macia, A. (1982): Generalización en el entrenamiento de estrategias cognitivas. En R. Fernández-Ballesteros (ed.), *Evaluación de contextos*, págs. 259-276. Murcia: Servicio de Publicaciones de la Universidad.
- Fernández-Ballesteros, R.; Juan-Espinosa, M.; Colom, R. y Calero, M.D. (1997): Contextual and personal sources of individual differences in intelligence: Empirical results. En J. Kingma y W. Tomic (eds.), *Advances in Cognition and Educational Practice, IV*, págs. 221-274, London: JAI Press.

- Ferrara, R.A.; Brown, A.L. y Campione, J.C. (1986): Children's learning and transfer of inductive reasoning rules: Studies in proximal development. *Child Development*, 57, págs. 1087-1099.
- Ferreiro, E. (1997): *Alfabetización. Teoría y práctica*. México: Siglo XXI.
- Ferreiro, E. y Pontecorvo, C. (1996): Los límites entre las palabras. En E. Ferreiro; C. Pontecorvo; N. Ribeiro e I. García Hidalgo, *Caperucita Roja aprende a escribir*, págs. 43-86. Barcelona: Gedisa
- Ferreiro, E. y Teberosky, A. (1982): *Los sistemas de escritura en el desarrollo del niño*. México: Siglo XXI.
- Feuerstein, R. (1979): *The Dynamic Assessment of Retarded Performers*. Baltimore, Maryland: University Park Press.
- Feuerstein, R. (1990): The theory of structural cognitive modifiability. En B. Presseisen (comp.). *Learning and thinking styles: classroom interaction*. Washington, D.C.: National Education Association, págs. 68-134.
- Feuerstein, R. (1991): La teoría de la modificabilidad estructural cognitiva. *Educación Cognitiva I*, 31-75. Mira Editores.
- Feuerstein, R. (1998a): *El modelo de Experiencia de Aprendizaje Mediado (EAM)*. Conferencia presentada en el «International Workshop in Structural Cognitive Modifiability». Jersusalem, Israel.
- Feuerstein, R. (1998b): *La modificabilidad cognitiva*. Conferencia presentada en el 4º Congreso Mundial de Educación Infantil. Madrid, 1998.
- Feuerstein, R. y Jensen, M.R. (1980): Instrumental Enrichment: Theoretical basis, goals and instruments. *The Educational Forum*, 44, 401-423.
- Feuerstein, R.; Rand, Y.; Haywood, H.C.; Kyrarn, L. y Hoffman, M.B. (1995): *LPAD, Learning Propensity Assessment Device (Manual). New Experimental Version*, Jerusalem, Israel: Hadassah-WIZO-Canada Research Institute.
- Feuerstein, R.; Rand, Y. y Hoffman, M.B. (1979): *The Dynamic Assessment of Retarded Performers: The Learning Potential Assessment Device Theory, Instruments, and Techniques*. Baltimore, MD: University Park Press.
- Feuerstein, R.; Rand, Y.; Hoffman, M.B. y Miller, R. (1980): *Instrumental enrichment*. Baltimore: University Park Press.
- Feuerstein, R.; Rand, Y. y Rynders, J.E. (1988): *Don't accept me as I am: Helping «retarded» people to excel*. New York: Plenum Press.
- Filho, L. (1973): *Los Test ABC de verificación de la madurez necesaria para el aprendizaje de la lectura y escritura*. Buenos Aires: Kapelus
- Flavell, J.H. (1976): Metacognitive aspects of problem solving. En L.B. Resnick (ed.), *The Nature of intelligence*, págs. 231-235. Hillsdale, N.J.: Erlbaum.
- Flavell, J.H. (1981): Cognitive monitoring. En W.P. Dickson (ed.), *Children's oral communication skills*, págs. 35-60. New York: Academic Press.
- Flavell, J.H. (1992): Perspectives on perspective taking. En H. Beiling y P. Pufall (eds.), *Piaget's theory: Prospects and possibilities*, págs. 107-139. Mahwah, Nueva Jersey: Erlbaum.
- Flavell, J.H. (1993): *El desarrollo cognitivo*. Madrid: Aprendizaje Visor. (Original en inglés: *Cognitive development* (2ª ed.). Englewood Cliffs, N.J.: Prentice-Hall, 1985).
- Fletcher, J.M.; Francis, D.J.; Rourke, B.P.; Shaywitz, S.E. y Shaywitz, B.A. (1992): The validity of discrepancy-based definitions of reading disabilities. *Journal of Learning Disabilities*, 25, 555-561.
- Fontes, S. y Fontes, A.I. (2001): El control experimental. En *Diseños de investigación en psicología*. Madrid: UNED.
- Foulds, G.A. (1949): Variation in the intellectual activities of adults. *American Journal of Psychology*, 62, 238-246.
- Foulds, G.A. y Raven, J.C. (1950): An experimental survey with Progressive Matrices (1947) set Y and II. *British Journal of Educational Psychology*, XX (II), 55-62.
- Freixas, A. y Luque, A. (1998): ¿A favor de las niñas? Notas en el debate sobre la escuela coeducativa. *Cultura y educación*, 9, 51-62.

- Fuchs, L.S. y Deno, S.L. (1994): Must instructionally useful performance assessment be based in the curriculum? *Exceptional Children*, 61, 15-24.
- Fuchs, L.S. y Fuchs, D. (1996): Combining performance assessment and curriculum-based measurement to strengthen instructional planning. *Learning Disabilities Research & Practice*, 11, 183-192.
- Gal'perin, P. Ya (1966): Kucheniui ob interiorizatsii. *Voprosy Psikhologii*, 6, págs. 20-29.
- Garb, E. y Kozulin, A. (1998): *I think... therefore I read: A cognitive approach to English teaching*. Jerusalem: Acadamon.
- Garb, E. y Kozulin, A. (1999): *Dynamic assessment of EFL at the pre-academic centers*. Unpublished report.
- García Madruga, J.A.; Martín Cordero, J.; Luque, J.L. y Santamaría, C. (1995): *Comprensión y adquisición de conocimientos a partir de textos*. Madrid. Siglo XXI
- García Sánchez, J.N. (2001): *Dificultades de Aprendizaje e Intervención Psicopedagógica*. Barcelona. Ariel.
- García Sánchez, J.N. (2002): *Aplicaciones de intervención psicopedagógica*. Madrid. Pirámide.
- García Sánchez, J.N. y Marbán, J.M. (2003): El proceso de composición escrita en alumnos con DA y/o BR: estudio instruccional con énfasis en la planificación. *Infancia y Aprendizaje*, 26 (1), págs. 97-113.
- Gardner, H. (1983): *Frames of mind: The theory of multiple intelligences*. New York: Basic Books. (Trad. cast.: *Estructuras de la mente: la teoría de las inteligencias múltiples*. México: Fondo de Cultura Económica, 1994).
- Garner, R. (1987): *Metacognition and reading comprehension*. Norwood, New Jersey: Ablex Publishing Corporation.
- Garner, R. y Alexander, P. (1982): Strategic processing of text: An investigation of the effects on adults question-answering performance. *Journal of Educational Research*, 75, págs. 144-148.
- Garner, R. y Alexander, P.A. (1989): Metacognition: Answered and unanswered questions. *Educational Psychologist*, 24, págs. 143-158.
- Garner, R. y Anderson, J. (1982): Monitoring of understanding research: Inquiry directions, methodological dilemmas. *Journal of Experimental Education*, 50, 70-76.
- Garner, R.; Macready, G.B. y Wagoner, S. (1984): Readers acquisition of the components of the text-lookback strategy. *Journal of Educational Psychology*, 76, 300-309.
- Gilabert, R. y Vidal-Abarca, E. (1990): *Comprender para Aprender. Un proyecto de innovación educativa* (Memoria de investigación). Consellería de Cultura, Educación y Ciencia. Generalitat Valenciana. Valencia.
- Gómez, P.F.; Valero, J.; Buades, R. y Pérez, A. (1995): *Test de Habilidades Metalingüísticas (THM)*. Madrid: EOS.
- González, A. (1996): Mejoras en la calibración de la comprensión. *Estudios de Psicología*, 56, 113-126.
- González, A. (1997): Metamemoria y aprendizaje de textos. *Estudios de Psicología*, 58, 59-83.
- González, M.D. (1984): El diagnóstico precoz como medida de las dificultades de aprendizaje de la lectura. Validez del ABC de Filho y del Reversal Test. *Revista de Psicología General y Aplicada*, 139, págs. 59-73.
- González Portal, M.D. (1985): *Exploración de las Dificultades Individuales de Lectura, nivel 1 (EDIL 1)*. Madrid: TEA ediciones.
- González, M.J. y Romero, J.F. (2001): Intervención psicoeducativa en Comprensión Lectora. En: J.F. Romero y M.J. González, *Prácticas de Comprensión Lectora. Estrategias para el aprendizaje*. Madrid. Alianza Editorial.
- González, M.M. y Palacios, J. (1990): La Zona de Desarrollo Próximo como tarea de construcción. *Infancia y Aprendizaje*, 51-52, 99-122.
- Goodenow, C. (1992): Strengthening the links between educational psychology and the study of social context. *Educational Psychologist*, 27, 177-195.
- Goodman, K.S. (1967): Reading: A Psycholinguistic Guessing Game. *Journal of the Reading Specialist*, 4, 126-135.

- Goodman, K.S. (1982a): Reading: A Psycholinguistic Guessing Game. En E.V. Gollasch (ed.), *Language and literacy, vol. 1*, págs. 19-31. Boston: Routledge Kegan Paul.
- Goodman, K.S. (1982b): The reading process: Theory and practice. En E.V. Gollasch (ed.), *Language and literacy, vol. 1*, págs. 33-43. Boston: Routledge Kegan Paul.
- Gough, P.B. (1976): One second of reading. En H. Singer y R.B. Ruddell (eds.), *Theoretical Models and Processes of Reading*. New York: K.I.R.A.
- Grant, D.A. y Berg, E.A. (1948): A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem. *Journal of Experimental Psychology*, 38, 404-411.
- Greshman, F.M. y MacMillan, D.L. (1997): Social competence and affective characteristics of students with mild disabilities. *Review of Educational Research*, 67, 4, 377-415.
- Grigorenko, E.L. y Sternberg, R.J. (1998): Dynamic testing. *Psychological Bulletin*, 124, págs. 75-111.
- Grigorenko, E.L.; Sternberg, R.J. y Ehrman, M.E. (2000): A theory-based approach to the measurement of foreign language learning ability: The CANAL-FT theory and test. *The Modern Language Journal*, 84, págs. 390-405.
- Guthke, J. (1992): Learning test: The concept, main research findings, problems and trends. *Learning and Individual Differences*, 4, págs. 137-151.
- Guthke, J. (1993): Current trends in theories and testing of intelligence. En J.H.M. Hamers; K. Sijtsma y A.J.J.M. Ruijsenaars (comps.), *Learning potential testing*. Amsterdam: Swets and Zeitlinger, págs. 13-20.
- Guthke, J. y Beckmann, J. (2000): The learning test concept and its application in practice. En C. S. Lidz y J. G. Elliott (comps.), *Dynamic assessment: Prevailing models and applications*, Greenwich, CT, Elsevier-JAI, págs. 17-69.
- Guthke, J.; Beckman, J.F.; Stein, H.; Vahle, H. y Rittner, S. (1995): *Adaptive Computergestützte Intelligenz-Lerntestbatterie (ACIL) (The Adaptive Computer-Assisted Intelligence Learning Test Battery)*. Mödlingen: Schuhfried GmbH.
- Guthke, J. y Löffler, M. (1983): A diagnostic program (learning test) for the differential assessment of school failure in 1st grade pupils. En H.D. Rösler; J.P. Das y I. Wald (eds.), *Mental and Language retardation*, págs. 41-50. Berlin: Deutscher Verlag der Wissenschaften.
- Guthke, J. y Stein, H. (1996): Are learning test the better version of intelligence tests? *European Journal of Psychological Assessment*, 12, págs. 1-13.
- Guthke, J. y Wiedl, K.H. (1996): *Dynamisches Testen (Dynamic assessment)*. Gotinga, Hogrefe.
- Guthke, J. y Wingefeld, S. (1992): The Learning Test concept: Origins, state of art, and trends. En H.C. Haywood y D. Tzuriel (comps.) *Interactive testing*. New York: Springer-Verlag, págs. 64-93.
- Hacker, D. (1998): Self-regulated comprehension during normal reading. En D. Hacher, J. Dunlosky y A. Graesser, *Metacognition in educational theory and practice*, págs. 165-192, Hillsdale, N. J.: Lawrence Earlbaum Associates.
- Hambleton, R.K. (1983): *Applications of item response theory*. Vancouver, BC: Educational Research Institute of British Columbia.
- Hambleton, R.K.; Swaminathan, H. y Rogers, H.J. (1991): *Fundamentals of item response theory*. Newbury Park, CA, Sage.
- Hamers, J.H.M.; Pennings, A. y Guthke, J. (1994): Training-based assessment of school achievement. *Learning and Instruction*, 4, págs. 347-360.
- Hare, V.C. y Smith, D.C. (1982): Reading to remember: Studies of metacognitive reading skills in elementary school-aged children. *Journal of Educational Research*, 75, 157-164.
- Haywood, H.C. (1997): Interactive assessment. En R. Taylor (comp.), *Assessment of individuals with mental retardation*. San Diego: Singular Publishing Group.
- Haywood, H.C.; Brown, A.L. y Wingefeld, S. (1990): Dynamic approaches to psycho-educational assessment. *School Psychology Review*, 19, 411-422.
- Haywood, H.C. y Tzuriel, D. (comps.) (1992): *Interactive assessment*. New York: Springer-Verlag.

- Heesels, M.G.P. y Hamers, J.H.M. (1993): The learning potential test for ethnic minorities. En J.H.M. Hamers; K. Sijtsma y A.J.J.M. Ruijsenaars (comps.), *Learning potential testing*. Amsterdam: Swets and Zeitlinger, págs. 285-311.
- Henry, M.K. (1993): The role of decoding in reading research and instruction. *Reading and Writing: An Interdisciplinary Journal*, 5, 105-112.
- Hernández Martín, A. y Quintero Gallego, A. (2001): *Comprensión y Composición escrita. Estrategias de aprendizaje*. Madrid: Síntesis.
- Herranz Alonso, J. (1987): Comprensión lectora y aprendizaje. *Comunidad educativa*, 154, págs. 6-17.
- Hunt, E. (1974): Quote the Raven? Nevermore. En L.W. Gregg (ed.), *Knowledge and cognition*, págs. 129-157. New York: Wiley.
- Inizan, A. (1983): *Cuándo enseñar a leer*. Madrid: Visor-Aprendizaje.
- Jackson, M.D. y McClelland, J.L. (1979): Processing determinants of reading speed. *Journal of Experimental Psychology: General*, 108, 151-181
- Jastak, J.E. y Jastak, S.R. (1978): *The Wide Range Achievement Test*, edición revisada. Washington, DC: Jastak Associates.
- Jastak, S. y Wilkinson, G.S. (1984): *The Wide Range Achievement Test-Revised: Administration manual*. Washington, DC: Jastak Associates.
- Jeffrey, I. (1997): *The Cognitive Modifiability Battery-Assessment and Intervention: Clinical perspectives of a language therapist*. Paper presented at the 6th Conference of the International Association for Cognitive Education (IACE). Stellenbosch, South Africa.
- Jeffrey, I. y Tzuriel, D. (1999): *The Cognitive Modifiability Battery (CMB): Applications of a Dynamic Instrument in Speech Language Therapy*. Unpublished manuscript, The Variety Learning Center. Vancouver, BC, Canada.
- Jensen, M.R. (1990): Change models and some evidence for phases and their plasticity in cognitive structures. *International Journal of Cognitive Education*, 1, 5-16.
- Jensen, M.R. (1992): Principles of change models in school psychology and education. En J. Carlson (comp.), *Cognition and educational practice, vol. 1B*, págs. 47-72. Greenwich, CT: JAI.
- Jensen, M.R. (1998): *MindLadder projects: Student, school, family and community learning in the information age*, manuscrito inédito. International Center for Mediated Learning, American InterContinental University.
- Jensen, M.R. (2000): The MindLadder model: Using dynamic assessment to help students learn to assemble and use knowledge. En C.S. Lidz y J.G. Elliott (comps.), *Dynamic Assessment: Prevailing models and applications*, Greenwich, CT, Elsevier-JAI, págs. 187-227.
- Jensen, M.R.; Feuerstein, R.; Rand, Y.; Kaniel, S. y Tzuriel, D. (1988): Cultural difference and cultural deprivation. En R.M. Gupta y P. Coxhead (eds.), *Cultural diversity and learning efficiency*, págs. 64-88. London: McMillan.
- Jensen, M.R. y Singer, J.L. (1987): *Structural cognitive modifiability in low functioning adolescents: An evaluation of Instrumental Enrichment* (Report submitted to the State of Connecticut Department of Education, Bureau of Special Education and Pupil Personnel Services, Hartford, CT). New Haven, CT: Yale University, Department of Psychology.
- Jiménez, J.E.; Gregg, N. y Díaz, A. (2004): Evaluación de habilidades fonológicas y ortográficas en adolescentes con dislexia y adolescentes buenos lectores. *Infancia y Aprendizaje*, 27 (1), 63-84
- Jiménez González, J.E. y Ortiz González, M.R. (1998): *Conciencia fonológica y aprendizaje de la lectura*. Madrid: Síntesis.
- Jiménez, J.E. y Rodrigo, M. (1994): Is it true that the Differences in Reading Performance between students with and without LD cannot be explained by IQ? *Journal of Learning Disabilities*, 27, 155-163.
- Jiménez, J.E.; Rodrigo, M.; Ortiz, M.R.; Guzmán, R. (1999): Procedimientos de evaluación e intervención en el aprendizaje de la lectura y sus dificultades desde una perspectiva cognitiva. *Infancia y Aprendizaje*, 88, 107-122

- Johnson, E.S. (1978): Validation concept learning strategies. *Journal of Experimental Psychology: General*, 107, 237-267.
- Johnston, P.H. (1989): *La evaluación de la comprensión lectora. Un enfoque cognitivo*. Madrid: Aprendizaje-Visor.
- Jolibert, J. (1984): *Former des enfants lecteurs*. París: Hachette.
- Jolibert, J. (1995): Formar niños lectores/productores de textos. *Textos de didáctica de la Lengua y la Literatura*, 5, 81-92.
- Just, M.A. y Carpenter, P.A. (1987): *The psychology of reading and language comprehension*. Boston: Allyn and Bacon.
- Just, M.A. y Carpenter, P.A. (1992): A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, 99, 122-149.
- Kaniel, S. y Tzuriel, D. (1992): Mediated learning experience approach in the assessment and treatment of borderline psychotic adolescents. En H.C. Haywood y D. Tzuriel (comps.), *Interactive assessment*. New York: Springer-Verlag, págs. 399-418.
- Kaniel, S.; Tzuriel, D.; Feuerstein, R.; Ben-Shachar, N. y Eitan, T. (1991): Dynamic assessment, learning, and transfer abilities of Jewish Ethiopian immigrants to Israel. En R. Feuerstein, P.S. Klein y A. Tannenbaum (comps.), *Mediated learning experience*. Londres: Freund, págs. 179-209.
- Kar, B.C.; Dash, U.N.; J.P. Das y J. Carlson (1993): Two experiments on the dynamic testing of planning. *Learning and Individual Differences*, 5, págs. 13-29.
- Karmiloff-Smith, A. (1994): *Más allá de la modularidad. La Ciencia Cognitiva desde la perspectiva del Desarrollo*. Madrid: Alianza.
- Kent, R.N. y Foster, S.L. (1977): Direct observational procedures: Methodological issues in naturalistic settings. En A.R. Ciminero; K.S. Calhoun y H.E. Adams (eds.), *Handbook of Behavioral Assessment*, New York: Wiley.
- Kidron, R. (1989): *Basic tasks in Math: Diagnostic Didactic System (4th ed.)*. Tel Aviv: Nitzan.
- Kintsch, W. (1986): Learning from text. *Cognition and Instruction*, 3, 87-108.
- Kintsch, W. y Van Dijk, T.A. (1978): Towards a model of text comprehension and production. *Psychological Review*, 85, 363-394.
- Kirk, S.A.; McCarthy, J.J. y Kirk, W.D. (1928/2004): *Test Illinois de Aptitudes Psicolinguísticas*. Madrid: TEA ediciones. (Original en inglés: *The Illinois Test of Psycholinguistic Abilities*. University of Illinois Press).
- Klahr, D. (1992): Information processing approaches to cognitive development. En M. H. Bornstein y M. E. Lamb (eds.), *Developmental Psychology: an advanced textbook (3^o ed.)*. Hillsdale, N.J.: Earlbaum.
- Klauer, K.J. (1993): Evaluation einer Evaluation: Stellungnahme zum Beitrag von Hager and Hasselhorn. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 25, págs. 322-327.
- Klein, P. y Feuerstein, R. (1985): Environmental variables and cognitive development. Identification of the potent factors in adult-child interaction. En *The at-risk infant* ed. Sh. Harrel y N.J., Anastasiow. Baltimore y London: Paul Brookes Publishing Co.
- Klein, P.; Weider, S. y Greenspan, S. (1987): A theoretical overview and empirical study of MLE: prediction of preschool performance from mother-infant interaction patterns. *Infant Mental Health Journal*, 8, págs. 110-129.
- Kletzien, S. y Bednar, M. (1990): Dynamic assessment for at-risk readers. *Journal of Reading*, 528-533.
- Kozulin, A. (2000): *Instrumentos Psicológicos. La educación desde una perspectiva socio-cultural*. Barcelona: Paidós.
- Kozulin, A. y Falik, L. (1995): Dynamic cognitive assessment of the child. *Directions in Psychological Science*, 4, 192-196
- Kozulin, A. y Garb, E. (2001): *Dynamic Assessment of EFL Text Comprehension of At-Risk Students*. Paper presented at the 9th Conference of the European Association for Research on Learning and Instruction. Fribourg, Switzerland, 2001.
- Kubinger, K.D. (1986): Adaptive Intelligenzdiagnostik (Adaptive intelligence testing). *Diagnostica*, 32, 330-344.

- Kurtz, B.E. y Borkowski, J.G. (1987): Development of strategic skills in impulsive and reflective children: A longitudinal study of metacognition. *Journal of Experimental Child Psychology*, 43, 129-148.
- Laberge, D. y Samuels, S.J. (1976): Toward a model of automatic information processing in reading. En H. Singer y R.B. Ruddell (eds.), *Theoretical Models and Processes of Reading*. Newark, Del.: International Reading Association.
- Laboratory of Comparative Human Cognition (1982): Culture and Intelligence. En R.J. Sternberg (comp.), *Handbook of human intelligence*. New York: Cambridge University Press (Trad. cast. *Inteligencia humana*. Barcelona: Paidós, 1987), págs. 642-719.
- Lauchlan, F. y Elliott, J.G. (1997): *The Use of the Cognitive Modifiability Battery (CMB) as an Intervention Tool for Children With Complex Learning Difficulties*. Symposium and paper presented at the 6th Conference of the International Association for Cognitive Education (IACE), Stellenbosch, South Africa.
- Laughon, P. (1990): The dynamic testing of intelligence: A review of three approaches. *School Psychology Review*, 19, págs. 459-470.
- Lauth, G. y Wiedl, K.H. (1985): Zur Veraendbarkeit der Test-Leistung im CFT-20 durch Instruktions Intensivierung (Establishing change in test performance in the CFT-20 through intensifying instructions). *Diagnostica*, 31, 200-209.
- Lázaro Martínez, A. (1980): *Prueba de Comprensión lectora*. Madrid: TEA ediciones.
- LeGagnoux, G.; Michael, W.B.; Hocesvar, D. y Maxwell, V. (1990): Retest effects on standardized structure-of-intellect ability measures for a sample of elementary school children. *Educational and Psychological Measurement*, 50, págs. 475-492.
- León, J.A. y García Madruga, J.A. (1989): Comprensión de textos e instrucción. *Cuadernos de Pedagogía*, 169, 54-66.
- Leontiev, A. (1981): The problem of activity in psychology. En J.V. Wertsch (comp.), *The concept of activity in Soviet psychology*. Armonk, N.Y.: Sharpe, págs. 37-71.
- Levy, C. (1999): *The discriminant validity of the Application of Cognitive Functions Scale (ACFS): A performance comparison between typically developing and special needs preschool children*. Unpublished Master's thesis; Touro College, New York, USA.
- Lidz, C.S. (1991): *Practitioner's guide to dynamic assessment*. New York: Guilford.
- Lidz, C.S. (1995): Dynamic assessment and the legacy of L.S. Vygotski. *School Psychology International*, 16, págs. 143-153.
- Lidz, C.S. (2000): The application of Cognitive Functions Scale (ACFS): An example of curriculum-based dynamic assessment. En C.S. Lidz y J.G. Elliott (comps.), *Dynamic Assessment: Prevailing models and applications*, Greenwich, CT, Elsevier-JAI, págs. 407-439.
- Lidz, C.S. (comp.) (1987): *Dynamic assessment: An interactional approach to evaluating learning potential*. New York: Guilford Press.
- Lidz, C.S. y Elliott, J.G. (comps.) (2000): *Dynamic assessment: Prevailing models and Applications* (6). Greenwich, CT, Elsevier-JAI.
- Lidz, C.S. y Jepsen, R.H. (1997): *Application of Cognitive Functions Scale Administration Manual: Research version*. Obtainable from the senior author at Graduate School of Education and Psychology; Touro College, 27 West 23rd Street, Room 536, New York, NY 10010.
- Lidz, C.S.; Jepsen, R.H. y Miller, M.B. (1997): Relationships between cognitive processes and academic achievement: Application of a group dynamic assessment procedure with multiply handicapped adolescents. *Educational and Child Psychology*, 14, 56-67.
- Lidz, C.S. y Thomas, C. (1987): The Preschool Learning Testing Device: Extension of a static approach. En C.S. Lidz (comp.), *Dynamic assessment: An interactional approach to evaluating learning potential*. New York: Guilford Press, págs. 288-326.
- Lin, M. y Zabrusky, K. M. (1998): Calibration of Comprehension: research and implication for education and instruction. *Contemporary Educational Psychology*, 23, 345-391
- Little, T.; Das, J.P.; Carlson, J. y Yachimowicz, D. (1993): A hierarchical model of mental abilities. *Learning and Individual Differences*, 5, 219-240.

- Lizasoain Hernández, L. y Joaristi Olariaga, L. (2000): El análisis de datos en la evaluación de programas educativos. *Revista de Investigación Educativa*, 18 (2), 357-379.
- Loper, A.B. y Murphy, D.M. (1985): Cognitive self-regulatory training for underachieving children. En D.L. Forrest-Presley; G.E. Mackinnon y T.G. Waller (eds.), *Metacognition, cognition and human performance*. Nueva York. Academic Press.
- Lukas, J.F. y Santiago, K. (2004): *Evaluación educativa*. Madrid: Alianza editorial.
- Lundberg, I.; Frost, J. y Petersen, O. (1988): Effects of an extensive program for stimulating phonological awareness in preschool children. *Reading Research Quarterly*, 23, 263-284.
- Lundberg, I. y Høien, T. (1991): Initial enabling knowledge and skills in reading acquisition: print awareness and phonological segmentation. En D. Sawyer y B. Fox (eds.), *Phonological Awareness in Reading. The evolution of current perspective*. New York: Springer-Verlag.
- Luque, A.; Molina, A. y Navarro, J.J. (2000): *Educación la Tolerancia. Una propuesta de trabajo*. Sevilla: Diada editorial.
- Luque, A.; Ortega, R. y Cubero, R. (1997): Concepciones constructivistas y práctica escolar. En M.J. Rodrigo y J. Arny (comps.) (1997), *La construcción del conocimiento escolar*. Barcelona: Paidós.
- Luria, A.R. (1984): *El cerebro en acción*. Barcelona: Martínez Roca (Original en inglés: *The working brain*. New York: Basic Books, 1973).
- Luria, A.R. (1995): *Conciencia y Lenguaje*. Madrid: Aprendizaje Visor.
- Maldonado, A. y Sebastian, E. (1987): *La segmentación de palabras: un prerrequisito del aprendizaje de la lectura*. Boletín del ICE de la UAM, 9, 1-14
- Maldonado, A.; Sebastián, E. y Soto, P. (1992): *Retraso en lectura: evaluación y tratamiento educativo*. Madrid: Ediciones de la U.A.M.
- Male, D.B. (1996): Metamemorial functioning of children with moderate learning difficulties. *British Journal of Educational Psychology*, 66, 2, 145-157
- Mandler, J.M. (1978): A code in the node: the use of a story schema in retrieval. *Discourse processes*, 1, 14-35.
- Manis, F. y Morrison, F. (1985): Reading disability: A déficit in rule learning? En L. Siegel y F. Morrison (eds.), *Cognitive development in atypical children*, págs. 1-26, New York: Springer-Verlag.
- Marchesi, A. (1999): Los alumnos con escasa motivación para aprender. En A. Marchesi; C. Coll y J. Palacios (comps.), *Desarrollo Psicológico y Educación*, 3. *Trastornos del desarrollo y Necesidades Educativas Especiales* (págs. 183-208). Madrid: Alianza.
- Marín, M.A. (1994): Evaluación de la lectura. En F. Cabrera; T. Donoso y M^a A. Marín, *El proceso lector y su evaluación*. Barcelona: Laertes.
- Martí, E. (2000): Procesos cognitivos básicos y desarrollo intelectual entre los 6 años y la adolescencia. En C. Coll, J. Palacios y A. Marchesi (comps.), *Desarrollo Psicológico y Educación*. 1. *Psicología Evolutiva* (págs. 329-349). Madrid: Alianza.
- Martí, E.; Teberosky, A. y García-Milà, M. (1999): *Desarrollo de las notaciones como memoria externa*. Comunicación presentada en las I Jornadas sobre Memoria. Barcelona, 12-14 de Marzo.
- Martín, E. (1999): Enseñar a pensar a través del currículo. En C. Coll, J. Palacios y A. Marchesi (comps.), *Desarrollo Psicológico y Educación*. 3. *Trastornos del desarrollo y necesidades educativas especiales*. Madrid: Alianza.
- Martín, E. y Marchesi, A. (1990): Desarrollo metacognitivo y problemas de aprendizaje. En C. Coll, J. Palacios y A. Marchesi (comps.), *Desarrollo Psicológico y Educación III. Necesidades educativas especiales y aprendizaje escolar*. Madrid: Alianza.
- Martín, E. y Solè, I. (2001): El aprendizaje significativo y la teoría de la asimilación. En C. Coll; J. Palacios y A. Marchesi (comps.), *Desarrollo Psicológico y Educación*, 2. *Psicología de la educación escolar* (págs. 89-113). Madrid: Alianza.
- Martín Núñez, M.C. (1993): *Incidencia de las estrategias metacognitivas en el aprendizaje de las ciencias sociales: un estudio con alumnos de 4º y 5º de E.G.B.* Tesis Doctoral. Madrid: UNED.

- Maruny Curto, L.; Ministrál Morillo, M. y Miralles Teixidó, M. (2001): *Escribir y Leer. Materiales curriculares para la enseñanza y el aprendizaje del lenguaje escrito, de tres a ocho años*. Centro de Publicaciones del MEC y Edelvives.
- Mastropieri, M.A. y Scrugg, T.E. (1997): Best practice in promoting reading comprehension in student with learning difficulties. *Remedial and Special Education*, 18, 4, 197-213.
- Mateos, M.M. (1991): Un programa de instrucción en estrategias de supervisión de la comprensión lectora. *Infancia y Aprendizaje*, 56, 61-76
- Mateos, M.M. (2001): *Metacognición y Educación*. Buenos Aires: Aique
- Mauri, T. y Mirás, M. (1996): *L'avaluació en el centre escolar*. Barcelona: Graò/ICE UB.
- Mayor Zaragoza, F. (2002): Ciudadanía democrática. Reinventar la democracia, la cultura de paz, la formación cívica y el pluralismo. En F. Imbernon (coord.); J. Majo; M. Mayer; F. Mayor Zaragoza; R. Menchú y J.C. Tedesco: *Cinco ciudadanías para una nueva educación*. Barcelona: Graò.
- McDaniel; Hines; Waddill y Einstein (1994): What makes Folk Tales Unique: Content familiarity, Causal Structure, Scripts or Superstructures? *Journal of Experimental Psychology*, vol. 20 (1), págs. 169-184.
- Mckeown, M.G.; Beck, I.L.; Sinatra, G.M. y Loxterman, J.A. (1992): The contribution of prior Knowledge and coherent text to comprehension. *International Reading Association*, 27, 79-93.
- McNamara, D.S. y Kintsch, W. (1996): Learning from texts: effects of prior knowledge and text coherence. *Discourse processes*, 22, 247-288.
- Meichenbaum, D. (1985): Teaching thinking: A cognitive behavioral perspective. En S.F. Chipman; J.W. Segal y R. Glaser, *Thinking and learning skills, Vol. 2. Research and open Questions*. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Meichenbaum, D. (1986): Metacognitive Methods of instruction: Currents status and future prospects. En M. Schwebel y C.A. Maher. *Facilitating cognitive development: International perspectives, programs and practices*. Londres: The Haworth Press.
- Meichenbaum, D.; Burland, S.; Gruson, L. y Cameron, R. (1985): Metacognitive assessment. En S.R. Ussen (ed.), *The growth of reflection in children*. New York: Academic Press.
- Meilán, E.M. y Vieiro, P. (2002): Las ilustraciones como ayudas extratextuales que mejoran el recuerdo y el aprendizaje de textos expositivos. En *Cultura y Educación*, 14 (2), 187-207.
- Meloth, M.S. (1990): Changes in poor readers: Knowledge of cognition and the association of knowledge of cognition with regulation of cognition and reading comprehension. *Journal of Educational Psychology*, 82, 792-798.
- Meltzer, L.J.; Katzir, T.; Miller, L.; Reddy, R. y Roditi, B. (2004): Academic self-perceptions, effort, and strategy use in students with learning disabilities: Changes over time. *Learning Disabilities Research and Practice*, 19 (2), 99-108.
- Meltzer, L.J.; Reddy, R.; Pollica, L.; Roditi, B.; Sayer, J. y Theokas, C. (2004): Positive and negative self-perceptions: Is there a cyclical relationship between teachers' and students' perceptions of effort, strategy use, and academic performance? *Learning Disabilities Research and Practice*, 19 (1), 33-44.
- Mercer, N. (2000): *Words & Minds. How we use language to think together*. Londres: Routledge.
- Meyer, B.J.F. (1984): Text dimensions and cognitive processing. En H. Mandl; N. Stein y T. Trabasso (eds.), *Learning and comprehension of text*. Hillsdale, N. J.: Lawrence Erlbaum Associates.
- Meyer, B.J.F. (1985): Prose Analysis: Purposes, Procedures, and Problems. En B. K. Britton y J. B. Black (eds.), *Understanding Expository Text*. Hillsdale, N. J.: Lawrence Erlbaum Associates.
- Meyer, B.J.F.; Brandt, D.M. y Bluth, G.J. (1980): Use of top level structure in text: Key for reading comprehension of night-grade students. *Reading Research Quarterly*, XVI, 1, 72-102.
- Meyer, B.J.F. y Freedle, R.O. (1984): Effects of discourse type on recall. *American Educational Research Journal*, 21, 1, 121-143.

- Michalski, S. (1988): *Das Diagnostische Programm 'Begriffsanaloges Klassifizieren' – ein Lerntest für leistungsversagende Schüler erster Klassen [The diagnostic program 'Begriffsanaloges Klassifizieren' (dP BAK): A learning test for poor performers in first classes]* Unpublished doctoral dissertation, Karl-Marx-University, Leipzig.
- Miller, J. R. (1985): A Knowledge-Based Model of Prose Comprehension: Application of Expository Texts. En B. K. Britton y J. B. Black (Eds.): *Understanding Expository Text*. Hillsdale, N.J.: Lawrence Earlbaum Associates.
- Minick, N. (1987): Implications of Vygotski's theories of dynamic assessment. En C.S. Lidz (comp.), *Dynamic assessment: An interactional approach to evaluating learning potential*. New York: Guilford Press, págs. 116-140.
- Mirás, M. (2000): La escritura reflexiva. Aprender a escribir y aprender acerca de lo que se escribe. *Infancia y Aprendizaje*, 89, 65-80.
- Mirás, M. (2001): Afectos, emociones, atribuciones y expectativas: el sentido del aprendizaje escolar. En C. Coll; J. Palacios y A. Marchesi (comps.), *Desarrollo Psicológico y Educación, 2. Psicología de la educación escolar* (págs. 309-325). Madrid: Alianza.
- Mirás, M. y Solè, I. (1990): La evaluación del aprendizaje y la evaluación en el proceso de enseñanza y aprendizaje. En C. Coll; J. Palacios y A. Marchesi (comps.), *Desarrollo Psicológico y Educación, 2. Psicología de la Educación escolar*, págs. 419-431. Madrid: Alianza.
- Moll, L.C. (1989): *Combining zones for learning: an ethnographic approach*. Comunicación presentada en la III European Conference for Research on Learning and Instruction. Madrid, septiembre.
- Montanero, M. (2000): *La instrucción de estrategias de comprensión en el ámbito sociolingüístico del segundo ciclo de la E.S.O. Una propuesta didáctica para "enseñar a comprender"*. Tesis doctoral inédita. Universidad de Extremadura.
- Montanero, M. (2001): Metacompreensión y aprendizaje a partir de textos. *Cultura y Educación*, 13, 3, 317-328.
- Montanero, M.; Blázquez, F. y León, J.A. (2002): Enfoques de intervención psicopedagógica para la mejora de las capacidades de comprensión en la Educación Secundaria. *Infancia y Aprendizaje*, 25, 1, 37-52.
- Montero, D. (1993): *Evaluación de la conducta adaptativa en personas con discapacidades. Adaptación y validación del ICAP*. Bilbao: Mensajero.
- Mora, J. (1986): Enriquecimiento Instrumental: Comentarios a una dicha prometida. *Siglo Cero*, 106, 40-44.
- Mora, J. (1991): *Evaluación del programa Comprender y Transformar*. Tesis doctoral. Universidad de Sevilla, 1991.
- Mora, J. (1994): *Dificultades del Aprendizaje. Proyecto docente*. Sevilla: Universidad de Sevilla.
- Mora, J. (1997): *Trastornos del Aprendizaje Escolar*. Apuntes y comunicación personal del autor durante sesiones de clase y tutorías. Sevilla: Universidad de Sevilla.
- Mora, J. (1998): *El programa Comprender y Transformar. Bases, Metodología y Evaluación*. Sevilla: Repiso y Reguera.
- Mora, J. y Aguilera, A. (2000): *Atención a la diversidad en educación: dificultades en el aprendizaje del lenguaje, de las matemáticas y en la socialización*. Sevilla: Ediciones Kronos.
- Mora, J. y Mora-Merchán, J. (1995): *Matriz de conductas cognitivas enriquecidas*. Sevilla: Universidad de Sevilla.
- Mora, J. y Mora-Merchán, J. (1998): *Comprender y Transformar: Materiales para la educación primaria*. Sevilla: Reguera/Repiso.
- Mora, J. y Moreno, F.J. (1998): *Rehabilitación cognitiva en sujetos con daño cerebral: la versión L de «Comprender y Transformar»*. Informe de Investigación. Departamento de Psicología Evolutiva y de la Educación. Universidad de Sevilla.
- Mora, J.A. (1993): *Batería Evaluadora de las Habilidades Necesarias para el Aprendizaje de la Lectura y Escritura*. Madrid: TEA ediciones.

- Mora Merchán, J. y Mora, J. (2000): Trastornos del desarrollo emocional. En: J. Mora y A. Aguilera (coords.), *Atención a la diversidad en educación: dificultades en el aprendizaje del lenguaje, de las matemáticas y en la socialización*. Sevilla: Ediciones Kronos.
- Moreno Pérez, F.J. y García Gómez, I. (2000): Dificultades en el desarrollo de la comprensión lectora. En J. Mora y A. Aguilera (coords.), *Atención a la diversidad en educación: dificultades en el aprendizaje del lenguaje, de las matemáticas y en la socialización*. Sevilla: Ediciones Kronos.
- Moreno, F.J. y Mora, J. (2002): *Rehabilitación cognitiva en sujetos con daño cerebral. Evaluación de la versión L del programa «Comprender y Transformar»*. Tesis Doctoral. Universidad de Sevilla.
- Moreno, F.J.; Pastor, G. y Piñero, R. (1995): Rehabilitación cognitiva en sujetos con daño cerebral sobrevenido: Una experiencia práctica. En VARIOS, *Daño cerebral traumático, Neuropsicología y Calidad de vida* (págs. 425-434). Madrid: Mapfre.
- Moreno, M.C. (1996): *Las teorías y los métodos en la investigación sobre el desarrollo socio-personal*. Sevilla: Kronos.
- Myers, N. A. y Perlmutter, M. (1978): Memory en the years from two to five. En P.A. Ornstein (ed.), *Memory development in children* (págs. 191-214). Hillsdale, N.J.: Earlbaum.
- Naglieri, J. y Das, J.P. (1987): Construct and criterion related validity of planning, simultaneous and successive cognitive processing tasks. *Journal of Psychoeducational Assessment*, 4, 353-363.
- Naglieri, J.A. y Das, J.P. (1988): Planning-attention-simultaneous-successive (PASS): A model for testing. *School Psychology Review*, 19, págs. 423-458.
- Navarro Hidalgo, J.J. (2004): Constructivismo y desarrollo potencial. Implicaciones educativas. *Perspectiva CEP. Revista de los Centros del Profesorado de Andalucía*, 8, 137-157.
- Navarro Hidalgo, J.J. (2005): Lectura y Solidaridad. *Textos de Didáctica de la Lengua y de la Literatura*, 38, 91-104.
- Navarro Hidalgo, J.J.; Lama Ruiz, H. y Molina Bernáldez, A. (2001): *Evaluación dinámica de procesos implicados en la lectura. Un enfoque metacognitivo (Manual)*. Manuscrito no publicado. Universidad de Sevilla.
- Navarro Hidalgo, J.J. y Mora Roche, J. (2002): *Evaluación de metaconocimientos sobre la comprensión de textos*. Informe de Investigación, Departamento de Psicología Evolutiva y de la Educación. Universidad de Sevilla.
- Navarro Hidalgo, J.J. y Mora Roche, J. (2003a): Evaluación de metaconocimientos sobre la comprensión de textos: análisis de un instrumento de medida en un estudio descriptivo. *Apuntes de Psicología*, 21 (2), 193-209.
- Navarro Hidalgo, J.J. y Mora Roche, J. (2003b): *Evaluación dinámica de metaconocimientos sobre la comprensión de textos*. Manuscrito no publicado. Universidad de Sevilla.
- Newman, D., Griffin, P. y Cole, M. (1989): *The construction zone*. New York. Cabridge University Press. (trad. cast.: *La zona de construcción del conocimiento: trabajando por un cambio cognitivo en educación*. Madrid: Morata, 1998).
- Newman, F. y Holzman, L. (1993): *Lev Vygotski: Revolutionary scientist*. Londres: Routledge.
- Nisbett, J. y Shucksmith, J. (1987): *Estrategias de aprendizaje*. Madrid: Santillana.
- Nisbett, R.E. y Wilson, T.D. (1977): Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84, 231-259.
- Novak, J.D. (1998): *Conocimiento y aprendizaje. Los mapas conceptuales como herramientas facilitadoras para escuelas y empresas*. Madrid: Alianza [Publicación original en inglés, en 1998].
- Nunziati, G. (1990): Pour construire un dispositif d'évaluation formative. *Cahiers Pédagogiques*, 280, 47-64.
- Olshavsky, J.E. (1976): Reading as problem solving: An investigation of strategies. *Reading Research Quarterly*, 12, 654-675.
- Olson, D.R. (1998): *El mundo sobre el papel*. Barcelona: Gedisa.
- Ontoria, A. (1993): *Mapas conceptuales. Una técnica para aprender*. Madrid: Narcea.
- Ortar, G. y Ben-Schachar, N. (1972): *Reading Comprehension Tests*. Jerusalem: Ministry of Education and Culture.

- Ortiz, M. (1994): *Conciencia Fonológica y Conciencia del Lenguaje escrito en el aprendizaje de la lectura*. Tesis Doctoral. Universidad de la Laguna.
- Ortiz, M^a R. y Jiménez, J.E. (1993): *Prueba de Conocimiento sobre el Lenguaje Escrito (CLE)*. Madrid: TEA ediciones.
- Palinscar, A.S. y Brown, A.L. (1984): Reciprocal teaching of comprehension. Fostering and metacognitive strategies. *Cognition and Instruction, 1*, 117-175.
- Palmer, J.; Mcleod, C.M.; Hunt, E. y Davidson, J.E. (1985): Information processing correlates of reading. *Journal of memory and language, 24*, 59-88.
- Paris, S.G.; Cross, D.R. y Lipson, M.Y. (1984): Informed strategies for learning: A program to improve children's reading awareness and comprehension. *Journal of Educational Psychology, 76*, 1239-1252.
- Paris, S.C.; Lipson, M.Y. y Wixson, K.K. (1983): Becoming a Strategic Reader. *Contemporary Educational Psychology, 8*, 293-316.
- Parrilla, M.A. (1996): *Apoyo a la escuela: un proceso de colaboración*. Bilbao: Ediciones Mensajero.
- Parrilla, A. y Daniels, H. (1998): *Creación y desarrollo de grupos de apoyo entre profesores*. Bilbao: Ediciones Mensajero.
- Pascual, G. y Goikoetxea, E. (2003): Resumen y formulación de preguntas: efectos sobre la comprensión lectora en niños de Primaria. *Infancia y Aprendizaje, 26* (4), págs. 439-450.
- Perfetti, C.A. y Hogaboam, T. (1975): The relationship between single word decoding and reading comprehension skill. *Journal of Educational Psychology, 67*, 461-469
- Peronard, M.; Crespo, N. y Guerrero, I. (2001): La Metacompreensión en escolares chilenos de Educación Básica. *Signos, XXXIV*, 49-50, 149-164.
- Peronard, M.; Crespo, N. y Velásquez, M. (2000): La Evaluación del Conocimiento Metacompreensivo en alumnos de Educación Básica. *Signos, XXXIII*, 47, 161-180.
- Peronard, M.; Crespo, N.; Velásquez, M. y Viramonte, M. (2002): Conocimiento Metacognitivo del Lenguaje escrito: instrumento de medida y fundamentación teórica. *Infancia y Aprendizaje, 25*, 2, 131-145.
- Piaget, J. (1969): *El nacimiento de la inteligencia en el niño*. Madrid: Aguilar (Original en francés: *La naissance de l'intelligence chez l'enfant*. Neuchâtel-París: Delachaux-Niestlé).
- Powell, A.; Love, L. y Vega, M. (1976): Further note on rule utilization in solution of abstract picture-analogy problems. *Psychological Report, 39*, 387-390.
- Pozo, J.I. (1996): *Aprendices y maestros. La nueva cultura del aprendizaje*. Madrid: Alianza editorial.
- Pozo, J.I.; Monereo, C. y Castelló, M. (2001): El uso estratégico del conocimiento. En C. Coll; J. Palacios y A. Marchesi (comps.), *Desarrollo Psicológico y Educación, 2. Psicología de la Educación escolar*, págs. 211-229. Madrid: Alianza.
- Pozo, J.I. y Postigo, Y. (1993): Las estrategias de aprendizaje como contenido del currículo. En C. Monereo (ed.), *Estrategias de aprendizaje: procesos contenidos e interacción*. Barcelona: Doménech.
- Ramos, J.L. y Cuetos, F. (2000): *Evaluación de los procesos lectores (PROLEC-SE)*. Madrid: TEA ediciones.
- Rasch, G. (1980): *Probabilistic models for some intelligence and attainment tests* (Copenhagen, Danmarks Paedagogiske Institut, 1960), Chicago: University of Chicago Press.
- Raven, J.C. (1995): *Test de Matrices Progresivas. Escala general y Escala color*. Madrid: TEA ediciones (Original en inglés: *Progressive Matrices, Sets A, B, C, D, & E*. Londres: H.K. Lewis & Co., 1938).
- Reder, L.M. (1996): Different Research Programs on Metacognition: Are the Boundaries Imaginary? *Learning and Individual Differences, 8*, 4, 383-390.
- Reid, M.K. y Borkowski, J.G. (1987): Causal attributions of hyperactive children: Implications for teaching strategies and self-control. *Journal of Educational Psychology, 79*, 296-307.
- Reinhardt, B.M. (1989): *Cognitive Modifiability in developmentally delayed children*. Unpublished doctoral dissertation. Yeshiva University, Ferkauf Graduate School of Psychology, New York.

- Repetto, E. (1997): Entrenamiento metacognitivo y ganancias significativas de los alumnos en algunas variables cognitivas y pedagógicas. *Revista española de pedagogía*, 206, 5-11.
- Repetto, E.; Beltrán, S.G.; Manzano, N. y Téllez, J.A. (2001): Evaluación del programa Comprender y Aprender en el aula. *Revista Electrónica de Investigación y Evaluación Educativa (RELIEVE)*, vol. 7, nº 24.
- Repetto, E.; Manzano, N.; Sutil, I.; Beltrán, S.G. y Téllez, J.A. (2001): *Elaboración de un cuestionario para evaluar las estrategias de metacomprensión lectora*. Comunicación presentada en el X Congreso Nacional de Modelos de Investigación Educativa. A Coruña: AIDIPE, págs. 260-264.
- Repetto, E.; Sutil, I.; Manzano, N. y col. (Téllez, J.A. y Beltrán, S.G.) (2002): *Comprender y Aprender en el Aula. Programa para la integración de las estrategias de lectura en el currículum*. Madrid: UNED (3 manuales del profesor y 3 cuadernos del alumno).
- Repetto Talavera, E.; Téllez Muñoz, J.A. y Beltrán Campos, S.G. (2002): *Intervención psicopedagógica para la mejora de la comprensión lectora y del aprendizaje*. Madrid: UNED
- Resing, W.C.M. (1990): *Intelligentie en leerpotentieel. Een onderzoek naar het leerpotentieel van jonge leerlingen uit het basis- en special onderwijs* [Intelligence and learning potential. Research on the learning potential of young children from mainstream and special education]. Lisse: Swets & Zeitlinger.
- Resing, W.C.M. (1993): Measuring inductive reasoning skills: The construction of a learning potential test. En J.H.M. Hamers; K. Sijtsma y A.J.J.M. Ruijsenaars (eds.), *Learning potential assessment: Theoretical, methodological and practical Issues*, págs. 219-242. Lisse/Berwyn PA: Swets & Zeitlinger.
- Resing, W.C.M. (1997): Learning potential assessment: The alternative for measuring intelligence? *Educational and Child Psychology*, 14, 68-82.
- Resing, W.C.M. (1998): Intelligence and learning potential. Theoretical and research issues. En W. Tomic y J. Kingma (eds.), *Advances in cognition and educational practice, vol. 5: Conceptual issues in research on intelligence*. Greenwich, CT: JAI Press, Inc.
- Resing, W.C.M. (2000): Assessing the learning potential for inductive reasoning (LIR) in young children. En C.S. Lidz y J.G. Elliott (comps.), *Dynamic assessment: Prevailing models and applications*. Greenwich, CT: Elsevier-JAI, págs. 224-262.
- Resing, W.C.M.; Slenders, A.P.; Bosma, A.C. y Iedema, R. (1997): *Assessing inductive reasoning skills in ethnic minority groups*. Paper presented at the 7th Conference of the European Association for Research on Learning and Instruction. Athens, Greece.
- Resing, W.C.M. y Van Wijk, A.M. (1996): Leepotentieel: Onderzoek bij allochtone leerlingen uit het basisonderwijs [learning potential: Research with ethnic minority children from mainstream education]. *Pedagogische Studiën*, 35, 432-444.
- Richgels, D.J.; Mgee, L.M.; Lomax, R.G. y Sheard, C. (1987): Awareness of four text structures: Effects on recall of expository text. *Reading Research Quarterly*, XXII, 2, págs. 177-196.
- Rivière, A. (1985): *La psicología de Vygotski*. Madrid: Aprendizaje Visor.
- Rivière, A. (1997): El tratamiento del autismo como trastorno del desarrollo: principios generales. En A. Rivière y J. Martos (comps.), *El tratamiento del autismo. Nuevas perspectivas*. Madrid. Ministerio de Trabajo y Asuntos Sociales.
- Rogoff, B. (1993): Pensamiento compartido y participación guiada: Conclusiones y especulaciones. En B. Rogoff, *Aprendices del pensamiento. El desarrollo cognitivo en el contexto social*, págs. 241-263. Barcelona: Paidós.
- Rogoff, B. y Wertsch, J. (comps.) (1984): *Children's learning in the zone of proximal development*. San Francisco: Jossey Bass.
- Romero, J.F. (1990): Las relaciones sociales en los niños con dificultades de aprendizaje. En: A. Marchesi; C. Coll y J. Palacios (comps.), *Desarrollo Psicológico y Educación III. Necesidades educativas especiales y aprendizaje escolar*. Madrid: Alianza Editorial.
- Romero, J.F. (1999): Retrasos madurativos y dificultades en el aprendizaje. En C. Coll, J. Palacios y A. Marchesi (comps.), *Desarrollo Psicológico y Educación. 3. Trastornos del desarrollo y necesidades educativas especiales* (págs. 75-96). Madrid: Alianza.

- Romero, J.F. y González, M.J. (2001): Estrategias de Comprensión Lectora. En: J.F. Romero y M.J. González: *Prácticas de Comprensión Lectora. Estrategias para el aprendizaje*. Madrid. Alianza Editorial.
- Romero, J.F. y Lavigne, R. (2005): *Dificultades en el Aprendizaje: Unificación de Criterios Diagnósticos. Definición, Características y Tipos*. Materiales para la práctica orientadora. Volumen I. Consejería de Educación. Junta de Andalucía. Dirección General de Participación y Solidaridad Educativa.
<http://www.juntadeandalucia.es/educacion/proa>
- Romero, J.F. y Lavigne, R. (2006): *Dificultades en el Aprendizaje: Unificación de Criterios Diagnósticos. Procedimientos de Evaluación y Diagnóstico*. Materiales para la práctica orientadora. Volumen II. Consejería de Educación. Junta de Andalucía. Dirección General de Participación y Solidaridad Educativa.
<http://www.juntadeandalucia.es/educacion/proa>
- Rosenblatt, L.M. (1978): *The reader, the text, the poem: The transactional theory of the literary work*. Carbondale, Illinois: Southern Illinois University Press.
- Rosenshine, B. y Meister, C. (1994): Reciprocal Teaching. A review of the Research. *Review of Educational Research*, 64 (4), 479-530.
- Rosenthal, R. (1966, 1976): *Experimenter Effects in Behavioral Research*. New York: Appleton-Century-Crofts.
- Ruddell, R.B. y Ruddell, M.R. (1994): Language acquisition and literacy processes. En R.B. Ruddell; M.R. Ruddell y H. Singer (eds.), *Theoretical models and processes of reading* (4ª ed., págs. 83-103). Newark, DE: International Reading Association.
- Ruddell, R.B. y Speaker, R.B. (1985): The interactive Reading Process: A Model. En H. Singer y R.B. Ruddell, *Theoretical Models and processes of reading*. Newark, Delaware: International Reading Association.
- Rueda, M.; Sánchez, E. y González, L. (1990): El análisis de la palabra como instrumento para la rehabilitación de la dislexia. *Infancia y Aprendizaje*, 49, 39-52
- Ruijsenaars, A.J.J.M.; Castelijns, J.H.M. y Hamers, J.H.M. (1993): The validity of learning potential tests. En J.H.M. Hamers; K. Sijtsma y A.J.J.M. Ruijsenaars (comps.), *Learning potential testing*. Amsterdam: Swets and Zeitlinger, págs. 69-82.
- Rumelhart, D. (1976): *Toward an Interactive Model of Reading. Technical Report, nº 56*. San Diego, California: Center for Human Information Processing. University of California.
- Rumelhart, D. (1977): Toward an Interactive model of reading. En S.D. Dornic (ed.), *Attention and Performance VI*. Hillsdale, N.J.: Erlbaum.
- Rumelhart, D. (1980): Schemata: The building blocks of cognition. En R.C. Spiro; B.C. Bruce y W.F. Brewer (eds.), *Theoretical issues in reading comprehension*. Hillsdale, N.J.: Erlbaum.
- Rumelhart, D. y McClelland, J.L. (1981): Interactive processing through spreading activation. En A.M. Lesgold y C.A. Perfetti (eds.), *Interactive processes in reading*, págs. 37-60, Mahwah, New Jersey: Erlbaum
- Saldaña Sage, D. (2001): *Evaluación dinámica de procesos metacognitivos*. Tesis doctoral. Universidad de Sevilla.
- Salomon, G. y Perkins, D.N. (1998): Individual and social aspects of learning. *Review of Research in Education*, 23, 1-24.
- Sánchez, E. (1989): *Procedimientos para instruir en la comprensión de textos*. Madrid: Ministerio de Educación y Ciencia, C.I.D.E.
- Sánchez, E. (1990): El aprendizaje de la lectura y sus problemas. En A. Marchesi; C. Coll y J. Palacios (comps.), *Desarrollo psicológico y educación III. Necesidades educativas especiales y aprendizaje escolar*, págs. 121-137. Madrid: Alianza Editorial.
- Sánchez, E. (1993): *Los textos expositivos. Estrategias para mejorar su comprensión*. Madrid: Aula XXI-Santillana.
- Sánchez, E. (1998): *Comprensión y Redacción de textos*. Barcelona: Edebe.

- Sánchez, E. (1999): El lenguaje escrito y sus dificultades: una visión integradora. En A. Marchesi; C. Coll y J. Palacios (comps.), *Desarrollo psicológico y educación, 3. Trastornos del desarrollo y necesidades educativas especiales*, págs. 127-155. Madrid: Alianza Editorial.
- Sánchez, E.; Orrantia, J. y Rosales, J. (1992): ¿Cómo mejorar la comprensión de textos en el aula? *Comunicación, Lenguaje y Educación, 14*, 89-112.
- Sánchez, E.; Rosales, J. y Suárez, S. (1999): Interacción profesor/alumnos y comprensión de textos. Qué se hace y qué se puede hacer. *Cultura y Educación, 14-15*, págs. 71-89
- Sarriá Sánchez, E. (2001): Métodos y diseños de investigación. En S. Fontes; C. García; A.J. Garriga; M^a C. Pérez-LLantada y E. Sarriá, *Diseños de investigación en Psicología*, págs. 95-122. Madrid: UNED.
- Saxe, G.B. (1977): A developmental analysis of notational counting. *Child Development, 48*, 1512-1520.
- Saxe, G.B. (1981): Body parts as numerals: A developmental analysis of numeration among the Oksapmin in Papua, N. Guinea. *Child Development, 52*, 306-316.
- Scardamalia, M. y Bereiter, C. (1984): Development of strategies in text processing. En H. Mandl; N. L. Stein y T. Trabasso (eds.), *Learning and comprehension of text*. Hillsdale, N.J.: Lawrence Earlbaum Associates.
- Schacter, D.L. y Tulving, E. (1994): What are the memory systems of 1994? En D. I. Schacter y E. Tulving (eds.), *Memory Systems 1994* (págs. 1-38). Cambridge, MA: The MIT Press.
- Schlee, J. (1985): Förderdiagnostik – Eine bessere Konzeption? En R.S. Jäger; R. Horn y K. Ingenkamp (comps.), *Tests and trends*, Weinheim, Beltz, págs. 82-208.
- Schmidt, L. (1971): Testing-the-limits in Leistungserhalten: Möglichkeiten and Grenzen. En E. Duhm (comp.), *Praxis der Klinischen Psychologie, vol. 2*, Gotinga, Hogrefe.
- Schöttke, H.; Bartram, M. y Wiedl, K.H. (1993): Psychometric implications of learning potential testing: A typological approach. En J.H.M. Hamers; K. Sijtsma y A.J.J.M. Ruijssenaars (comps.), *Learning potential testing*. Amsterdam: Swets and Zeitlinger, págs. 153-173.
- Scrugg, T.E. y Mastropieri, M.A. (1992): Classroom applications of mnemonic instruction: Acquisition, maintenance and generalization. *Exceptional Children, 58*, 219-231.
- Serpell, R. (1993): *The significance of schooling: Life journeys in an African society*. New York: Cambridge University Press.
- Sewell, T. W. y Price, V. D. (1991): MLE: implications for achievement motivation and cognitive performance in low socioeconomic and minority children. En R. Feuerstein; P. Klein y A. Tennenbaum (comps.), *Mediated learning experience*. Londres: Freund.
- Shurin, R. (1998): *Validity and reliability of the Application of Cognitive Functions Scale with preschool children with disabilities*. Unpublished masters thesis; Touro College, New York, N.Y. ERIC document TM 030312.
- Sigel, I. (1982): The relationship between parental distancing strategies and the child's cognitive behavior. En L. M. Laosa e I. E. Sigel (eds.), *Families as learning environments for children*. Nueva York: Plenum Press.
- Sijtsma, K. (1993): Classical and modern test theory with an eye toward learning potential testing. En J.H.M. Hamers; K. Sijtsma y A.J.J.M. Ruijssenaars (comps.), *Learning potential testing*. Amsterdam: Swets and Zeitlinger, págs. 117-133.
- Sinatra, G. M.; Beck, I. L. y McKeown, M. G. (1993): How Knowledge influenced two intervention designed to improve comprehension. *Reading Psychology: An International Quarterly, 14*, 141-163.
- Sinclair, A.; Jarvella, R. y Levelt, W. (1978): *The child's conception of language*. Berlin. Springer-Verlag.
- Slenders, A.P.A.C. y Resing, W.C.M. (1997): Leerpotentieelonderzoek bij allochtone leerlingen: een verkenning met de Leerpotentieeltest voor Inductief Redeneren [Research on learning potential with children of ethnic minority groups: the Learning Potential Test for Inductive Reasoning]. *Pedagogische Studiën, 74*, 183-196.
- Smith, F. (1983): *Comprensión de la lectura. Análisis psicolingüístico de la lectura y su aprendizaje*. México: Trillas.

- Solè, I. (1987): Las posibilidades de un modelo teórico para la enseñanza de la comprensión lectora. *Infancia y Aprendizaje*, 39-40, págs. 1-13
- Solè, I. (1992): *Estrategias de lectura*. Barcelona: Graó/ICE
- Solè, I. (2001): Leer, lectura, comprensión: ¿hemos hablado siempre de lo mismo? En Varios Autores, *Comprensión lectora. El uso de la Lengua como procedimiento*. Caracas, Edit. Laboratorio Educativo/ Barcelona, Graó.
- Solè, I. y Teberosky, A. (2001): La enseñanza y el aprendizaje de la alfabetización: una perspectiva psicológica. En C. Coll; J. Palacios y A. Marchesi (comps.), *Desarrollo Psicológico y Educación, 2. Psicología de la educación escolar*. Págs. 461-485. Madrid: Alianza.
- Stahl, S. y Murray, B. (1994): Defining Phonological Awareness and Its Relationship to Early Reading. *Journal of Educational Psychology*, 86, 221-234
- Stanovich, K.E.; Cunningham, A.E. y Cramer, B.R. (1984): Assessing phonological awareness in kindergarten children: issues of task comparability. *Journal of Experimental Child Psychology*, 38, 175-190
- Stein, N.L. y Nezworski, M.T. (1978): The effect of organization and instructional set on story memory. *Discourse processes*, 1, 177-193.
- Sternberg, R.J. (1985): *Beyond IQ: A triarchic theory of human intelligence*. New York: Cambridge University Press (Trad. cast.: *Más allá del cociente intelectual*. Bilbao: Desclée de Brouwer, 1990).
- Sternberg, R.J. (1997a): *Successful intelligence*. New York: Plume (Trad. cast.: *Inteligencia exitosa: cómo una inteligencia práctica y creativa determina el éxito en la vida*. Barcelona: Paidós, 1997).
- Sternberg, R.J. (1997b): The concept of intelligence and its role in lifelong learning and success. *American Psychologist*, 52, 1030-1037.
- Sternberg, R.J. (1998): Abilities are forms of developing expertise. *Educational Researcher*, 27, págs. 11-20.
- Sternberg, R.J. (1999a): Human intelligence: A case study of how more and more research can lead us to know less and less about psychological phenomenon, until finally we know much less than we did before we started doing research. En E. Tulving (comp.), *Memory, consciousness, and the brain: The Tallinn Conference*. Filadelfia, PA: Psychology Press, págs. 363-373.
- Sternberg, R.J. (1999b): Intelligence as developing expertise. *Contemporary Educational Psychology*, 24, 259-375.
- Sternberg, R.J.; Ferrari, M.; Clinkenbeard, P.R. y Grigorenko, E.L. (1996): Identification, instruction, and assessment of gifted children: A construct validation of a triarchic model. *Gifted Child Quarterly*, 40, 129-137.
- Sternberg, R. J. y Grigorenko, E. L. (2003): *Evaluación dinámica. Naturaleza y mediación del potencial de aprendizaje*. Barcelona: Paidós.
- Sternberg, R.J.; Grigorenko, E.L.; Ferrari, M. y Clinkenbeard, P.R. (1999): A triarchic analysis of an aptitude-treatment interaction. *European Journal of Psychological Assessment*, 15, págs. 1-11.
- Sternberg, R.J.; Grigorenko, E.L.; Ngorosho, D.; Tantufuye, E.; Mbise, A.; Nokes, C.; Jukes, M. y Bundy, D.A. (2001): Assessing Intelligence Potential in Rural Tanzania School Children. *Intelligence*, vol. 29, n° 5, págs. 401-409.
- Stevens, R. J. (1988): Effects of strategy training on the identification of the main idea of expository passages. *Journal of Educational Psychology*, 80, 1, 21-26.
- Suárez, A. y Meara, P. (1992): *CLT. Dos pruebas de comprensión lectora – Procedimiento «Cloze»*. Madrid: TEA ediciones.
- Swanson, H.L. (1992): Generality and modifiability of working memory among skilled and less skilled readers. *Journal of Educational Psychology*, 84, 473-488.
- Swanson, H.L. (1993): An information processing analysis of learning disabled children's problem solving. *American Educational Research Journal*, 30, 4, 861-893

- Swanson, H.L. (1995a): Effects of dynamic assessment on the classification of learning disabilities: The predictive and discriminant validity of the Swanson-Cognitive Processing Test (S-CPT). *Journal of Psychoeducational Assessment*, 13, 204-229.
- Swanson, H.L. (1995b): *Swanson-cognitive processing test*. Austin, TX: PRO-ED.
- Swanson, H.L. (1995c): Using the cognitive processing test to assess ability: Development of a dynamic assessment measure. *School Psychology Review*, 24 (4), 672-693.
- Swanson, H.L. (2000): Swanson-Cognitive Processing Test: Review and Applications. En C.S. Lidz y J.G. Elliott (comps.), *Dynamic Assessment: Prevailing Models and Applications*. Greenwich, CT: Elsevier-JAI, págs. 71-107.
- Swanson, H.L. y Trahan, M.F. (1992): Learning disabled reader's comprehension of computer mediated text: The influence of working memory, metacognition and attribution. *Learning Disabilities Research and Practice*, 7, 74-86.
- Swanson, H.L. y Trahan, M. (1996): Learning disabled and average reader's working memory and comprehension: does metacognition play a role? *British Journal of Educational Psychology*, 66, 3, 333-355
- Taylor, B. M. y Beach, R.W. (1984): The effects of text structures instruction on middle-grade students' comprehension and production of expository text. *Reading Research Quarterly*, XIX, 2, 134-146.
- TEA (1971): *PMA. Test de Aptitudes Mentales Primarias*. Madrid: TEA ediciones.
- TEA (1976): *TIG-1. Test de Inteligencia General Dominos. Forma I*. Madrid: TEA ediciones.
- Téllez, J.A. (2004): *La comprensión de los textos escritos. Evaluación del programa Comprender y Aprender en el aula*. Tesis Doctoral. UNED.
- Terman, L.M. (1970): *Concept Mastery Test*. New York: Psychological Corporation.
- Tolchinsky, L. (1990): La reproducción de relatos en niños entre cinco y siete años: Organización sintáctica y funciones narrativas. *Anuario de Psicología*, 47, 64-87.
- Toro, J. y Cervera, M. (1984): *Test de Análisis de lecto-escritura (TALE)*. Madrid: Aprendizaje Visor.
- Tulviste, P. (1991): *The cultural-historical development of verbal thinking*. Nueva York: Nova Scientia Publishers.
- Tunmer, W. y Fletcher, C. (1981): The relationship between conceptual tempo, phonological awareness, and word recognition in beginning readers. *Journal of Reading Behavior*, 13, 173-186
- Tunmer, W. y Herriman, M. (1984): The Development of Metalinguistic awareness: A conceptual overview. En W.E. Tunmer; C. Pratt y M.L. Herriman (eds.), *Metalinguistic Awareness in Children*. Berlin: Springer-Verlag.
- Tzuriel, D. (1992): The dynamic testing approach: A reply to Frisby and Braden. *The Journal of Special Education*, 26, 302-324.
- Tzuriel, D. (1995a): *Dynamic-interactive testing: The legacy of L.S. Vygotski and current developments*. Manuscrito inédito.
- Tzuriel, D. (1995b): *The Cognitive Modifiability Battery (CMB): Assessment and Intervention – Instruction Manual*. School of Education, Bar Ilan University.
- Tzuriel, D. (1997): *Symposium on the Cognitive Modifiability Battery (CMB): Assessment and Intervention*. Presentation at the 6th Conference of the International Association for Cognitive Education (IACE). Stellenbosch, South Africa.
- Tzuriel, D. (1999): *The Cognitive Modifiability Battery (CMB): Development of a novel dynamic assessment and intervention instrument for young children*. Unpublished manuscript, School of Education, Bar Ilan University.
- Tzuriel, D. (2000): The Cognitive Modifiability Battery (CMB): Assessment and intervention. En C.S. Lidz y J.G. Elliott (comps.), *Dynamic Assessment: Prevailing Models and Applications*. Greenwich, CT: Elsevier-JAI, págs. 375-406.
- Tzuriel, D. y Feuerstein, R. (1992): Dynamic group assessment for prescriptive teaching: Differential effect of treatment. En H.C. Haywood y D. Tzuriel (eds.), *Interactive Assessment*, págs. 187-206. New York: Springer-Verlag.

- Tzuriel, D. y Klein, P. (1985): Analogical thinking modifiability in disadvantaged, regular, special education and mentally retarded children. *Journal of Abnormal Child Psychology*, 13, 539-552.
- Tzuriel, D. y Klein, P. (1987): Assessing the young child: Children's analogical thinking modifiability. En C.S. Lidz (ed.), *Dynamic Assessment: An interactional approach to evaluating learning potential*, págs. 268-282. New York: Guilford Press.
- Tzuriel, D. y Samuels, M.T. (2000): Dynamic assessment of learning potential: Inter-rater reliability of deficient cognitive functions, types of mediation, and non-intellective factors. *Journal of Cognitive Education and Psychology*, n° 1, págs. 2-23.
- Valiña, Bernal y De Vega, (1990): Procesamiento estructural de historias: ¿reglas, macrorreglas o modelos mentales? *Revista de Psicología General y Aplicada*, n° 43, vol. 2, págs. 155-167
- Valmaseda, M. (1999): Los problemas de lenguaje en la escuela. En C. Coll; J. Palacios y A. Marchesi (comps.), *Desarrollo Psicológico y Educación*, 3. *Trastornos del desarrollo y necesidades educativas especiales* (págs. 101-123).
- Van Dijk, T.A. (1980): *Macrostructures*. Hillsdale, New Jersey: Erlbaum.
- Van Dijk, T.A. (1983): *La ciencia del texto*. Barcelona: Paidós.
- Van Dijk, T.A. y Kintsch, W. (1983): *Strategies of discourse comprehension*. New York: Academic Press.
- Varios Autores (1998): *Seminario T.A.E. Monográfico sobre lecto-escritura*. Texto sin publicar. Universidad de Sevilla.
- Vermunt, J.D. y Verloop, N. (1999): Congruence and friction between learning and teaching. *Learning and Instruction*, 9, 257-280.
- Vidal-Abarca, E. (1989): Estrategias para la comprensión y el recuerdo de la información más importante en textos expositivos: un estudio experimental. *Revista de Psicología de la Educación*, 2, 1-20.
- Vidal-Abarca, E. y Gilabert, R. (1991): *Comprender y Aprender*. Madrid: CEPE
- Vidal-Abarca, E. y Martínez Rico, G. (2001): ¿Por qué los textos son difíciles de comprender? Las inferencias son la respuesta. En Varios Autores, *Comprensión lectora. El uso de la Lengua como procedimiento*. Caracas, Edit. Laboratorio Educativo/ Barcelona, Graó.
- Voss, J.F. (1982): *Knowledge and social science problem solving*. Nueva York: Aere.
- Vygotski, L.S. (1931/1997): La colectividad como factor de desarrollo del niño deficiente. En L. S. Vygotski, *Obras Escogidas V, Fundamentos de defectología*, (págs. 213-234). Madrid: Aprendizaje Visor.
- Vygotski, L.S. (1934/1995): *Pensamiento y Lenguaje*. Barcelona: Paidós.
- Vygotski, L.S. (1979/1996): *El desarrollo de los procesos psicológicos superiores*. Barcelona: Grijalbo Mondadori.
- Wagner, R.K.; Torgesen, J.K. y Rashotte, C.A. (1994): Development of reading-related phonological processing abilities: New evidence of bidirectional causality from a latent variable longitudinal study. *Developmental Psychology*, 30 (1), 73-87.
- Wertsch, J.V. (1988): *Vygotski y la formación social de la mente*. Barcelona: Paidós.
- Wertsch, J.V. (1991): A sociocultural approach to socially shared cognition. En L.B. Resnick; J.M. Levine y S.D. Teasley (comps.), *Perspectives on socially shared cognition*. Washington, DC: American Psychological Association, págs. 85-100.
- Wertsch, J.V. y Stone, C.A. (1985): The concept of internalization in Vygotski's account of the genesis of higher mental functions. En J.V. Wertsch (ed.), *Culture, communication and cognition*. Cambridge: Cambridge University Press.
- Wertsch, J.; Tulviste, P. y Hagstrom, F. (1993): A sociocultural approach to agency. En A. Forman; N. Minick y C. Addison Stone (comps.), *Contexts for learning*. Nueva York: Oxford University Press.
- Whitman, Th. (1990): Self-regulation and mental retardation. *American Journal on Mental Retardation*, 94, 4, 347-362.

- Wiedl, K.H. y Bethge, H.J. (1983): Die Anpassung der aufgabenbezogenen Betrachtungszeit an variierende Aufgabenschwierigkeiten: Deskriptive und Veraenderungsbezogene Analysen bei kognitiv impulsiven und reflexiven Kindern [The relationship of observation time to task demands: Analyses involving impulsive and reflexive children]. *Zeitschrift fuer Differentielle und Diagnostische Psychologie*, 4, 67-77.
- Wiedl, K.H. y Herring, D. (1978): Der Effekt der Situationsbezogenheit schulischer leistungsbeurteilung auf die Vorhersagegueltigkeit des «Farbigen-Matrizen-Lerntests» und des CFT-1-Grundintelligenztests. *Zietschrift fuer Psychologie*, 186, 471-476.
- Winograd, P.N. (1984): Strategic difficulties in summarizing texts. *Reading Research Quarterly*, 19, 404-425. Traducción en *Infancia y Aprendizaje*, 31-32, 67-87.
- Wong, B.Y.L. (1985): Metacognition and learning disabilities. En D.L. Forrest-Presley; G.E. Mackimon y T.G. Waller (eds.), *Metacognition, cognition and human performance*. Vol. 2. *Institutional practices*. Nueva York: Academic Press.
- Wong, B.Y.L. (1986): Metacognition and special education: a review of a view. *The Journal of Special Education*, 20, 9-29.
- Wong, B.Y.L. (1994): Instructional parameters promoting transfer of learned strategies in students with learning dissabilities. *Learning Dissabilities Quarterly*, 17, 110-120.
- Yela, M. (1967): El factor espacial en la estructura de la inteligencia técnica. *Revista de Psicología General y Aplicada*, XXII, 609-635.