

FACULTAD DE CIENCIAS DE LA EDUCACIÓN

UNIVERSIDAD DE SEVILLA



Trabajo de Fin de Grado

Grado en Pedagogía

**UN ESTUDIO SOBRE RENDIMIENTO ESCOLAR:
SELECCIÓN DE CENTROS ESCOLARES DE MUY ALTA Y
MUY BAJA EFICACIA EN LA COMUNIDAD AUTONÓMICA
ANDALUZA.**

AUTOR/A: Carla Ortiz de Villate Fernández

TUTOR: Javier Rodríguez Santero

CURSO: 2018/2019

ÍNDICE

Resumen	1
1. Introducción	3
2. Fundamentación teórica	6
2.1. Un estudio sobre eficacia escolar: las pruebas de evaluación educativa gran escala.....	6
2.2. Antecedentes y estado del estudio sobre eficacia y rendimiento escolar.....	7
2.3. Definición de escuela eficaz: aproximación al concepto de ganancia residual y valor añadido.....	11
2.4. Los Modelos Jerárquicos Lineales.....	16
3. Objetivos del estudio	22
3.1. Objetivos generales.....	22
3.2. Objetivos específicos.....	22
4. Método	23
4.1. Muestra participante.....	23
4.2. Instrumentos y variables.....	26
4.3. Procedimiento.....	31
4.4. Análisis de datos.....	37
5. Resultados	46
5.1. Covariables significativas en el estudio.....	47
5.2. Selección de centros.....	51
6. Conclusiones	54
7. Limitaciones	56
8. Referencias bibliográficas	59

RESUMEN

El objetivo del presente trabajo es identificar y seleccionar los centros escolares de muy alta (CAEF) y muy baja eficacia (CBEF) en la Comunidad Autónoma Andaluza, considerando un centro de alta eficacia aquel que presente un alto valor añadido una vez sean controlados los efectos de las variables contextuales (residuo). De la misma forma, se considera CBEF, aquellos centros que presentan un valor añadido muy inferior al esperable. Para ello, se tomarán las muestras de las puntuaciones obtenidas en razonamiento matemático y comunicación lingüística correspondientes a las evaluaciones diagnósticas (ED) realizadas por la Agencia Andaluza de Evaluación Educativa (AGAEVE), en los cursos 2013/14, 2014/15 y 2016/17.

Como criterios para el análisis y selección de centros más y menos eficaces, se tendrán en cuenta el promedio de sus residuos y los residuos extremos, es decir, aquellos que muestren una tendencia creciente (o decreciente) en el modelo longitudinal contextualizado. Para ello, se emplearán técnicas estadísticas multivariante, concretamente, modelos jerárquicos lineales (HLM), ya que estos respetan la estructura anidada de los datos y nos permite analizar de forma conjunta los efectos de las variables presentes en el contexto educativo a diferentes niveles. Como resultado del estudio, de los 1936 centros analizados, se seleccionaron finalmente 50 CAEF y 50 CBEF haciendo referencia a los criterios: promedio de los residuos y residuos extremos.

PALABRAS CLAVE: *Eficacia del centro de enseñanza, Valor añadido, Mejora escolar, Modelos jerárquicos lineales, Evaluaciones diagnósticas, Estudio longitudinal, Estudio contextualizado.*

ABSTRACT

The aim of this work is to identify and select the highly effective schools (CAEF) and the low effective schools (CBEF) in the Autonomous Community of Andalusia, considering a school as highly effective when presents a high added value once the effects of contextual variables have been controlled (residual). In the same way, schools are considered as low effective when presents an added value lower than expected. To get it, samples will be taken of the results obtained in mathematical reasoning and linguistic communication corresponding to the diagnostic evaluations (ED) by the Andalusian Agency for Educational Evaluation (AGAEVE), in 2013/14, 2014/15 and 2016/17.

As criterion for the analysis and selection of the most and the least efficient schools, we will use the average of their residues and extreme residues, those show an increasing (or decreasing) trend in the contextualized longitudinal model. In order to identify the schools which have the best and the worst results, multivariate statistical techniques will be used, specifically, hierarchical linear modeling (HLM). These techniques respect the nested structure of the data and allow us to jointly analyze the effects of the variables present in the educational context at different levels. As a result of the study, we analyzed 1936 schools and finally we have selected 50 CAEF and 50 CBEF using the following criterion: extreme scores and residual.

KEYWORDS: *School Effectiveness, Value Added Models, School improvement, Hierarchical linear modeling, Diagnosis evaluation, Longitudinal studies, contextualized studies.*

1. INTRODUCCIÓN

La finalidad de este trabajo es identificar y seleccionar aquellos centros escolares que podrían considerarse de muy alta o muy baja eficacia escolar en la Comunidad Autónoma Andaluza. Para ello, se utilizarán los datos proporcionados por la Agencia Andaluza de Evaluación Educativa (AGAEVE). Esta, es una de las agencias a nivel nacional que lleva a cabo pruebas de evaluación educativas. Concretamente, las denominadas Pruebas ESCALA. El objetivo de su aplicación, no es más que el de dar a conocer tanto a los integrantes de los centros educativos que participan en las evaluaciones, como a la propia Administración Educativa, cuál es el estado del sistema educativo en el momento en el que se realizan.

Actualmente, el fenómeno de la evaluación de los sistemas educativos ha estado más vigente, debido a la necesidad de dar explicación a multitud de factores relacionados con la eficacia escolar. Entre ellos, evaluar si las prácticas educativas que se llevan a cabo en los centros son eficaces para alcanzar una formación integral en su alumnado. Además, las aplicaciones de las evaluaciones diagnósticas (ED) ayudan a establecer un dictamen acerca de la posición, en función del rendimiento de sus alumnos, en las que encuentra cada centro y, a nivel general por Comunidades Autónomas. Este es un tema muy controvertido, puesto que, la consecución de unos buenos resultados no depende únicamente de aquellos factores que el centro educativo puede controlar. En este contexto, sería erróneo constatar que un centro eficaz es aquel que obtiene las mejores puntuaciones en las competencias evaluadas, ya que habría que medir la influencia de multitud de elementos que lo caracterizan y contemplar la complejidad de la realidad educativa.

La aplicación de las evaluaciones educativas de carácter censal han sido un instrumento clave en la aparición y consolidación de los estudios sobre eficacia y mejora escolar. Gracias a los datos de carácter contextual y, por supuesto, los resultados de las pruebas recogidos por las ED, se ha intentado dar una respuesta lógica a la complejidad educativa que se nos presenta. La existencia de múltiples variables contextuales, por parte del propio alumno y del centro educativo hacen que, inevitablemente, influyan sobre su aprendizaje. Es importante tener consciencia de que la consecución de unos buenos

resultados y, por ende, un buen rendimiento, no es una función que se pueda dejar en manos del centro educativo solamente. En esta situación, entran en juego la función de las familias y el propio alumnado, teniendo en cuenta la influencia externa que este tiene.

Desgraciadamente, es casi imposible realizar una labor de control total de todos los factores contextuales que pueden influir en el rendimiento. Aun así, en los últimos años, en España, se están llevando a cabo estudios que intentan identificar estos elementos y, retraer el efecto de aquellos que son puramente contextuales y, en los cuales, el centro o el alumnado no tienen ningún poder de decisión. De esta forma, y en relación a los estudios de eficacia y mejora escolar, se plantea como último objetivo, proporcionar un catálogo de buenas prácticas para ayudar a los centros a conseguir buenos resultados en sus alumnos y desarrollar sus capacidades de la mejor forma posible. En conclusión, favorecer la calidad de la educación.

Como se ha comentado anteriormente, el estudio aquí presente, está centrado en la Comunidad Autónoma Andaluza (CAA). Concretamente, se han empleado los resultados de las pruebas ESCALA administradas en 2º de Educación Primaria durante los cursos 2013/14, 2014/15 y 2016/17 por la AGAEVE. Las pruebas de evaluación ESCALA permiten obtener información objetiva y rigurosa sobre las destrezas de escritura, cálculo y lectura, que se adquieren en los primeros años de escolaridad obligatoria y son básicas, ya que determinarán en buena medida la evolución escolar del alumnado. Este es un instrumento que permite conocer, no solo las competencias del alumnado, sino la situación de los centros y del Sistema Educativo Andaluz en su conjunto.

En las evaluaciones educativas realizadas por la AGAEVE se encuentran asociados unos cuestionarios de contexto que normalmente son rellenados por las familias de los alumnos que van a realizarlas. Su objetivo es dar a conocer el carácter socioeconómico y cultural de cada centro y de los alumnos de forma individual, entre otras funciones. Estos cuestionarios de familia, serán los que nos proporcionarán los datos relevantes acerca de las variables contextuales que puedan afectar al rendimiento final y que, consecuentemente, introduciremos en los modelos de estudio. Así, se intentará, de la mejor forma posible, realizar un control de los efectos que estas producen y ajustar las

puntuaciones de los centros a lo que cabría esperar de ellos, considerando su contexto (residuos).

En este trabajo se realizará un análisis cuantitativo basado en técnicas estadísticas multinivel (HLM), con el objetivo de identificar y seleccionar aquellos centros de muy alta y muy baja eficacia escolar. Se efectuará, por un lado, un estudio transversal contextualizado, considerando las cortes transversales en las cohortes de 2º de primaria, para las competencias de razonamiento matemático y comunicación lingüística. Por otro lado, se realizará un estudio longitudinal, en el que se analizará el crecimiento o decrecimiento de los residuos de los centros en los tres años seleccionados. Para ambos estudios, como se mencionó anteriormente, se utilizarán modelos jerárquicos lineales, ya que son las únicas técnicas estadísticas que, hasta el momento, son capaces de responder a la estructura multinivel de los datos que presenta la realidad educativa.

Posteriormente, en un estudio distinto a este, se pretende seguir con un análisis de carácter más cualitativo, para conocer las prácticas educativas que emplean los centros seleccionados y diseñar un catálogo de buenas prácticas orientado a la mejora escolar. Sin embargo, esto último, no forma parte de este trabajo, tratándose así solo de una orientación hacia la línea que se pretende seguir en futuras investigaciones.

2. FUNDAMENTACIÓN TEÓRICA

2.1. Un estudio sobre eficacia escolar: las pruebas de evaluación educativa a gran escala.

La evaluación de los sistemas educativos es un fenómeno que ha cobrado importancia en los últimos años debido a la necesidad de explicar y controlar los factores que influyen en los resultados de sus estudiantes y, consecuentemente, de los propios centros educativos. Básicamente, las entidades encargadas de esta evaluación tienen la función de comprobar y registrar los resultados procedentes del rendimiento de los alumnos en las diferentes etapas educativas. De esta forma, comprobar el efecto que estas causan socialmente, proporcionar una visión de la situación actual del sistema educativo y, ofrecer una retroalimentación. El objetivo final de esta retroalimentación, no tiene que ser otro que el de la optimización y mejora de la calidad de la educación.

Se conoce que, en la evaluación educativa, la aplicación de las pruebas de diagnóstico en España difiere en función de la Comunidad Autónoma en la que se empleen. Esto, en cierta parte, puede estar relacionado con la descentralización en política educativa que existe. Aun así, la aplicación de las evaluaciones diagnósticas proporciona un dictamen acerca de la posición, en función del rendimiento de sus alumnos, en las que encuentra cada centro y, a nivel general por CCAA. No se debe tratar, de ninguna manera, de establecer un ranking de posicionamiento, sino de un procedimiento de control acerca de los centros educativos que desempeñan un mejor funcionamiento y, posteriormente, aplicar sus prácticas en los centros que tienen mayores dificultades.

En general, y según Martínez-Abad, Chaparro, Lizasoain, Caso y Urías (2014), este tipo de pruebas surgieron con la finalidad de dar a conocer las capacidades y conocimientos de los alumnos en los diferentes estadios educativos y ofrecer una guía de prácticas efectivas que ayuden a mejorar la educación. De igual modo, y según Antillón, Larrazolo y Backhoff (2006), el principal objetivo de las evaluaciones educativas a gran escala, es el de ofrecer la información general acerca del rendimiento académico de los estudiantes evaluados, ya sea para realizar diagnósticos, selecciones o mejoras educativas. Generalmente, el diseño de las pruebas de evaluación se lleva a cabo mediante

su estandarización, siempre en función de los objetivos y metas que se planteen en cada una de estas (Ántillon et al., 2006).

Las puntuaciones obtenidas por estas asociaciones en las pruebas ayudan en la toma de decisiones educativas, ya sea a nivel de centro, individuo o sistema educativo en general (Ántillon et al., 2006). La evaluación no sería útil si, después de conocer los resultados de los centros, no se diera a conocer sus resultados y los aspectos a mejorar. En España, es necesario poner en marcha este tipo de iniciativas que asistan y respalden la verdadera función de la evaluación: enriquecer y potenciar la calidad de la educación basándose en las prácticas de aquellos centros que obtienen un muy alto valor añadido y aplicándolas en los centros que hayan obtenido un rendimiento desfavorable (bajo valor añadido) en los resultados de las evaluaciones (Martínez-Abad et al., 2014).

2.2. Antecedentes y estado del estudio sobre eficacia y rendimiento escolar.

El campo de investigación en el que se enmarca el estudio de la eficacia y la mejora escolar es conocido por el acrónimo SESI (School Effectiveness and School Improvement). El movimiento de la mejora de la eficacia escolar es una vertiente de investigación de carácter teórico-práctica que busca ampliar el conocimiento existente acerca de los procesos de transformación educativa, propulsando cambios reales en los centros escolares que, de alguna forma, abalen el incremento del rendimiento de los estudiantes (Mac-Gilchrist, Myers y Reed, 2004).

Existen dos corrientes de investigación que sustentan este movimiento. Una de ellas, podríamos decir, de carácter más cuantitativa, es el estudio de la eficacia escolar basado en el rendimiento de los alumnos pertenecientes a los centros escolares y, por otro lado, la investigación enfocada hacia la mejora escolar, de carácter más cualitativo. Ambas se fusionan en este movimiento, ya que buscan el reconocimiento de aquellos factores que influyen positivamente en el aprendizaje y, por lo tanto, en el rendimiento de los alumnos, con el fin de una mejora de la calidad de la educación (Mata y Ballesteros, 2012).

El movimiento de la eficacia escolar tiene como objetivo reconocer aquellas causas que intervienen de forma eficaz en el rendimiento del alumnado y que, de esta forma, los centros educativos obtengan puntuaciones superiores a las esperadas teniendo en cuenta el contexto situacional donde se enmarca el estudio (Mata y Ballesteros, 2012). Se puede decir, por lo tanto, que este tipo de estudios marcan la relación entre la eficacia escolar y la mejora de los centros educativos porque, a pesar de poder considerarse corrientes distintas, guardan una estrecha relación.

El estudio de la eficacia escolar tiene un extenso recorrido en el campo de la investigación educativa. Esta comenzó en el año 1966 con el Informe Coleman, en el cual se observó la necesidad de incrementar el conocimiento e información sobre el rendimiento escolar puesto que, era un ámbito poco estudiado hasta el momento e investigaciones relacionadas ayudaron a conocer la situación de la enseñanza y mejorar la calidad educativa (Gamazo, Martínez-Abad, Olmos-Migueláñez y Rodríguez-Conde, 2018). Además, los resultados del Informe Coleman mostraban que mientras que los factores que hacían referencia al efecto de las escuelas no tenían un peso considerable en los resultados, los factores que conciernen al contexto sociocultural de los alumnos, si lo tenían. Siendo así, se empezó a observar la necesidad de considerar en los estudios de rendimiento y eficacia escolar, aquellas variables contextuales que repercuten en el alumno, ya que son determinantes para la lograr resultados reales y válidos en los estudios (Mata y Ballesteros, 2012).

Según Mata y Ballesteros (2012), otras investigaciones realizadas proporcionaron avances metodológicos relevantes, entre las que podemos destacar, aquellas que dieron lugar a los modelos multiniveles, que eran capaz de dar una respuesta más real a la situación propuesta. Un ejemplo lo tenemos en los estudios realizados durante los años setenta y ochenta. Podríamos destacar, tal y como afirman Gamazo et al. (2018):

Los estudios de Weber, que añadió variables procesuales como el clima escolar o el liderazgo a un objeto de estudio que hasta entonces se centraba en factores contextuales, los trabajos de Brookover, Beady, Flood, Schewitzer y Wisenbaker,

quienes ampliaron la investigación de la influencia de los procesos escolares con el empleo de grandes muestras, o los hallazgos de Aitkin y Longford, a partir del uso de los modelos estadísticos multinivel, realizando contribuciones clave que ayudaron a avanzar en el estudio de la eficacia escolar, estableciendo las bases de lo que hoy es un campo consolidado de investigación en educación. (p. 58).

Los trabajos de Reynolds, Hopkins y Stoll (1993) o los de Creemers y Reezigt (2005) podrían considerarse también unos de los primeros realizados en este campo, aunque quizás la obra más relevante sea el *International Handbook of School Effectiveness and Improvement* (Townsend y Avalos, 2007) que supone una excelente síntesis del estado de la cuestión sobre esta línea de investigación. Otro antecedente en cuanto a este tipo de investigaciones lo tenemos en el estudio de Creemers, Stoll y Reezigt (2007), quienes realizaron un análisis internacional comparativo acerca de la eficacia y mejora escolar, proponiendo la combinación de estos para llevar a cabo un marco de referencia a la mejora de la eficacia escolar en los centros educativos.

Según Gamazo et al. (2018), en los últimos años, el estudio y análisis de la eficacia escolar ha cobrado importancia alrededor del mundo. Los estudios sobre eficacia y rendimiento escolar nos han ayudado a conocer cuáles son las prácticas educativas que resultan eficaces en los centros y cómo llevarlas a cabo (Hernández-Castilla, Murillo y Martínez-Garrido, 2013). La realización de pruebas de evaluación educativas tanto a nivel internacional como de tipo nacional se han incrementado con el objetivo de ofrecernos una mayor seguridad y control del sistema educativo y, en conclusión, conseguir transformar la realidad educativa y favorecer la obtención de resultados positivos en todos sus ámbitos (Martínez-Abad, Lizasoain, Castro y Joaristi, 2017).

Este campo de investigación educativa es una línea de trabajo que ha ido, como se ha mencionado anteriormente, incrementando su desarrollo, uniendo a su vez, el estudio de la eficacia con la posible mejora educativa. Es decir, si se obtienen los centros considerados de muy alta eficacia y se analizan en profundidad cuáles son las prácticas educativas que llevan a cabo, se podrían obtener resultados fiables acerca de cómo mejorar la calidad de la educación.

Según Gamazo, Olmos-Migueláñez y Martínez (2016), los estudios relacionados con el rendimiento y la mejora escolar han ido aumentando, además por la variedad y accesibilidad a las bases de datos con origen en las pruebas de evaluación educativas a gran escala que realizan las agencias tanto nacionales como internacionales.

Las pruebas de diagnóstico no solo nos aportan datos relevantes acerca del rendimiento que tienen los alumnos en las diferentes competencias evaluadas, sino que, además, nos aportan una gran cantidad de referencias sobre aquellas variables contextuales que pueden influir en el aprendizaje de los alumnos y el rendimiento de los centros educativos. Gracias a estos datos, podemos realizar estudios contrastados para conocer cuáles son los factores que indiquen mayormente en el rendimiento final del alumno y, consecuentemente, en las propias escuelas (De la Orden y Jornet, 2012).

Uno de los criterios que más se manejan en el campo de la investigación educativa para estimar la eficiencia de los sistemas educativos en general, es el rendimiento académico obtenido por el alumnado (Iñiguez-Berrozpe y Marcaletti, 2018). Estas pruebas pueden servir como base para la realización de estudios de rendimiento escolar, puesto que, como se ha mencionado anteriormente, nos ofrecen información acerca de variables contextuales que pueden afectar directa o indirectamente a los resultados de las evaluaciones. Según Gamazo et al. (2018), gracias a la variedad de estudios existentes acerca del rendimiento escolar del alumnado y a sus variables contextuales como son el nivel socioeconómico del centro y del alumnado, el nivel de implicación familiar y las características demográficas y organizativas del centro, entre otras, han ido aumentando el número de investigaciones destinadas al análisis de estas para, en un futuro, mejorar la calidad educativa.

Por último, sería importante comentar brevemente el estado de la cuestión en nuestro país, España, y qué investigaciones son pioneras en este ámbito de investigación. González (2015) señala que, en nuestro país, el desarrollo y puesta en marcha del movimiento de la eficacia y mejora escolar ha sido muy limitado e insuficiente para explicar los fenómenos que se dan y aparecen en la realidad educativa. Esta situación ha ocasionado que, a pesar de los intentos por innovar en este tipo de investigaciones, se conozcan un número muy reducido de estudios de carácter empírico que se dediquen a analizar y a investigar acerca de la eficacia y la mejora escolar (Murillo, 2000).

De entre ellos, podemos destacar los estudios realizados a finales de los años setenta por Manuel Millán Ventura, el cual realiza una revisión de las investigaciones que hasta el momento se habían llevado a cabo y hace hincapié en aquellas que han obtuvieron resultados exitosos (González, 2015).

También podemos destacar, una década más tarde, el estudio realizado por Aurora Fuentes Vicente llamado *Procesos funcionales y eficacia de la escuela*. Gracias a este estudio, la autora nos aporta un perfeccionamiento en las técnicas utilizadas para la investigación educativa puesto que, ofrece una visión empírica que ayuda al reconocimiento de aquellas variables propias de la eficacia escolar (González, 2015). Además, encontramos las aportaciones de Javier murillo (2004) y Luis Lizasoain en (2014), quienes han dado una visión más actual sobre la investigación de la eficacia escolar en España. En esta última, se ofreció a la comunidad educativa, además, una caracterización de los centros seleccionados como de alta eficacia, indicando así que, factores como el Índice Socioeconómico y Cultural familiar (ISEC), el género del estudiante, la inmigración que se da en los centros y el tipo de modelo educativo que se desempeñe, afectan al rendimiento final de alumnos y centros educativos en el País Vasco.

2.3. Definición de escuela eficaz: aproximación al concepto de ganancia residual y valor añadido

La investigación sobre eficacia escolar ha obtenido un papel muy importante tanto en el ámbito académico como en el social, ya que han aportado conocimientos útiles para la toma de decisiones en educación (González, 2015). Murillo (2005), está de acuerdo con esta afirmación, ya que considera la investigación de la eficacia escolar como una de las más influyentes en materia educativa.

Para conocer el fenómeno de la eficacia escolar más en profundidad, sería importante realizar una aproximación al concepto de eficacia o escuela eficaz. Las primeras definiciones de eficacia en el ámbito educativo hacían referencia a los conceptos de calidad y equidad, puesto que la evaluación de la calidad del rendimiento del alumnado

nos aporta una perspectiva real acerca del funcionamiento de los centros escolares y de cuáles son los que obtienen los mejores y peores resultados. Estos, a su vez, se podrían comparar, para llegar a deducciones sobre las pautas de actuación que funcionan en el ámbito educativo (Cuttance, 1992). Sin embargo, no basta con conocer qué centros han obtenido los mejores resultados, sino que, también es necesario conocer cuáles son los centros educativos que han alcanzado los mejores resultados, siempre y cuando se hallan controlando los efectos contextuales que influyen notablemente en el proceso. Es decir, compensando las desigualdades, y alcanzando puntuaciones académicas superiores a las que se esperaban (valor agregado) y, teniendo en cuenta su realidad social (Murillo, 2005).

De esta forma, según Murillo (2005), se puede llegar a considerar una escuela eficaz como aquella que obtiene un alto rendimiento en sus alumnos y, por consiguiente, un desarrollo integral de todos estos, habiendo controlado en todo momento la situación previa de sus integrantes y las variables socio-económicas y culturales que intervienen en el contexto educativo. Básicamente, escuelas eficaces son los centros que obtienen resultados superiores a los esperados. De la misma forma, las escuelas de baja eficacia son aquellas en las que los estudiantes obtienen resultados por debajo de lo que se podría esperar considerando el efecto del rendimiento previo y las variables contextuales que afectan tanto al alumno como al centro (Hernández-Castilla et al., 2013). La definición de escuela eficaz realizada por Murillo tiene especial relación con la que Joaristi, Lizasoain y Azpillaga (2014) añaden. En esta, ellos consideran escuela eficaz como aquella que obtienen, en las materias analizadas, una puntuación media superior a la que se esperaba (alto valor añadido) considerando la influencia de las variables contextuales.

Por otro lado, en otros estudios, los centros de alta eficacia o escuelas más eficaces se han definido como los centros que obtienen los mejores resultados, teniendo en cuenta lo que les cabría esperar de ellos y, relacionándolos con los que más ganancias o valor añadido tienen. Siguiendo con la misma definición, una escuela ineficaz, sería aquella que obtiene resultados inferiores a los previstos, por lo que ganan menos (Mata y Ballesteros, 2012).

Esta definición, en mi opinión, es un poco pobre. No solo se debe puntualizar la importancia de las puntuaciones para ser considerado como un centro de alta eficacia, sino que, un centro eficaz debe ser aquel que, además de obtener las mejores puntuaciones durante un periodo considerable en el tiempo, también es capaz de conseguir un desarrollo integral del alumnado en general. Los alumnos deben de adquirir una serie de competencias y habilidades en todos los ámbitos ya sea, de tipo instrumental en las diferencias materias, como a nivel social, ético y emocional. Es aquí donde entra en juego el papel de la mejora educativa y la investigación a nivel cualitativo para conocer los aspectos y prácticas que verdaderamente hacen a una escuela eficaz.

Comentar que, en el análisis de centros alta eficacia, la eficacia diferencial se relaciona con las ganancias residuales, es decir, la diferencia que existe entre la puntuación media obtenida por los centros educativos en las competencias evaluadas y la puntuación estimada o prevista en función de las variables contextuales que influyen en el modelo y, por lo tanto, en el rendimiento final del alumnado (Martínez-Abad et al., 2017). De esta forma, la definición más correcta de centros de alta eficacia (CAEF), y que más se corresponde con los objetivos del estudio, se refiere a aquellos centros que obtengan altos residuos y positivos en los modelos jerárquicos lineales de carácter contextualizado, es decir, que sus puntuaciones medias obtenidas sean muy altas en comparación con las esperadas una vez que la influencia externa de las variables que influyen en el rendimiento de los estudiantes haya sido controlada (Lizasoain, Bereziartua y Bartau, 2016). Esto se puede explicar, considerando que, como argumentan Lizasoain et al. (2016), un centro de alta eficacia no debe ser el que obtenga las mejores puntuaciones en las competencias evaluadas, ya que estaríamos hablando de puntuaciones brutas y en materia educativa, no representa la realidad.

La realidad educativa implica multitud de factores que hay que considerar para poder explicarla y, seleccionar centros en función de puntuaciones extremas sin un análisis previo de todas las variables contextuales que influyen sería un indicador claro de poca validez del estudio. Se debe realizar, para dar una respuesta certera, un ajuste de las puntuaciones, y para ello, es necesario modelizar mediante procedimientos

estadísticos de regresión multinivel y llevar el control de todos los factores que puedan incidir en el rendimiento del alumno. A través de los modelos jerárquicos lineales podríamos obtener la diferencia residual entre la puntuación media obtenida por el centro en las competencias seleccionadas y la puntuación media que se espera de este.

La diferencia residual nos ayuda a identificar los centros de alta y baja eficacia, ya que si un centro ha obtenido un residuo muy positivo en relación a lo que se cabría esperar en función de su contextualización y demás factores que puedan influir en el aprendizaje del alumnado, es que este está llevando a cabo unas prácticas que reflejan su calidad y eficacia como centro. De forma equivalente, si la diferencia residual que obtenemos en el análisis multinivel es muy negativa en comparación con las puntuaciones que esperábamos, indican claramente la ineficacia del centro. Esta ganancia o pérdida en función de la diferencia residual entre valor esperado y valor obtenido, hace referencia a medidas de valor añadido puesto que, los centros con altos residuos podrían considerarse como centros con altas ganancias o alto valor añadido (Martínez-Abad et al., 2017).

En todo momento se está haciendo referencia al término de valor añadido y este, es un término muy controvertido y que origina multitud de debates dentro de la evaluación de los sistemas educativos y los estudios del rendimiento y mejora escolar (Martínez-Abad, et al., 2014). Podemos encontrar el origen de este concepto en estudios e investigaciones de carácter económico para hacer referencia a cuestiones sobre el crecimiento o decrecimiento de bienes y riquezas. Entre las primeras aportaciones en materia educativa que incluyen el concepto de valor añadido, podemos destacar a Bryk y Weisberg, quienes en los años setenta presentaron su validez para indicar el crecimiento o decrecimiento de los centros educativos (Martínez-Abad, et al., 2014). Ellos muestran especial atención a la necesidad de utilización del término valor añadido para estimar el crecimiento esperado controlando los efectos de las variables contextuales y dando inicio a los primeros estudios sobre este tipo de enfoques aplicado a la investigación educativa.

Según Fitz-Gibbon (1997), actualmente, el uso del término valor añadido representa un progreso relativo, es decir, un progreso que está determinado por los diferentes contextos situacionales dónde se den. Consecuentemente, en investigación

educativa, valor añadido se refiere al control y aislamiento de aquellos factores de carácter escolar que inciden en el aprendizaje del alumnado, distinguiéndolos de aquellas variables individuales y sociales que, de una forma u otro, influyen en el proceso de evaluación (Martínez-Abad et al., 2017).

Más concretamente, Mata y Ballesteros (2012) consideran el valor añadido en el ámbito educativo como las prácticas, destrezas y metodologías de aprendizaje que un centro ofrece a su alumnado con el objetivo de contribuir a su desarrollo integral, una vez que se hayan controlado los efectos de las variables contextuales que influyen tanto a nivel de centro escolar, como a nivel de alumno. El valor añadido es, por tanto, un indicador de calidad y eficacia de un centro educativo. La idea fundamental es que los mejores resultados o, mejor dicho, los centros eficaces no se correspondan de forma directa con las puntuaciones obtenidas por los alumnos en las pruebas de evaluación, sino que se trate de un proceso de análisis más implacable y equitativo basados en los valores añadidos de los centros una vez que se tiene controlado el efecto que tienen las variables contextuales (Joaristi et al., 2014).

Aun así, para incrementar la validez de los resultados, sería aconsejable analizar las aportaciones del centro a nivel de rendimiento a lo largo de un periodo de tiempo, ya que si el mismo centro se ha mantenido en las primeras posiciones y su diferencia residual aumenta o se mantiene, es un centro de muy alta eficacia sin lugar a dudas. Asimismo, complementar la investigación con estudios de carácter longitudinal y que midan el crecimiento (o decrecimiento) de los residuos o puntuaciones en los años analizados, ayudaría a reforzar el valor de los resultados.

Por otro lado, y siguiendo con los argumentos que relacionan el valor añadido con las investigaciones de evaluación educativas, según González (2015), podemos señalar tres tipos de medidas que lo caracterizan. En primer lugar, recalcar el nivel de logro de los estudiantes considerando y controlando las variables sociales y culturales que puedan afectar en su rendimiento. En segundo lugar, señalar la función de valor añadido para identificar las ganancias de los propios estudiantes a partir del análisis de lo anteriormente

nombrado: resultados obtenido y variables contextuales que influyen en su aprendizaje. Por último y, en tercer lugar, identificar el valor agregado como progreso en red ajustado a todos sus antecedentes individuales tales como rendimiento previo, habilidades, competencias, etc.

Concluyendo, podemos decir que el valor añadido aplicado a la investigación educativa, y más concretamente a la evaluación de los sistemas educativos nos ofrece la oportunidad de establecer conclusiones reales, ya que opera desde lo que se espera que cada alumno y centro consiga y, en función de los contextos situacionales que les rodean (Aguado y Ballesteros, 2012). Por todo esto, es importante considerar todas estas medidas, y analizarlas mediante modelos jerarquizados multinivel, de lo cuales hablaremos más adelante, para establecer cuáles serán los centros de alta eficacia (CAEF) y centros de baja eficacia (CBEF) de la Comunidad Autónoma Andaluza.

2.4. Los Modelos Jerárquicos Lineales

Anteriormente, los datos de las evaluaciones educativas como PISA, en España, se analizaban a través de análisis basados en regresiones lineales o logísticas, pero este tipo de análisis no era el adecuado, ya que no se tenían en cuenta los posibles efectos que las variables contextuales ejercían entre los niveles de estudio (Blanco-Blanco, López-Martín y Ruiz de Miguel, 2014). Al obviar los efectos, no se prestaban atención a todas aquellas variables representativas al contexto que podían ser significativas en el estudio de rendimiento del alumnado, ya que estas, a su vez, no eran independientes y se tenían que considerar agrupadas a ambos niveles: alumnos y centros educativos. (Iñiguez-Berrozpe y Marcaletti, 2018).

La regresión lineal simple implica únicamente la relación entre dos variables, una de carácter dependiente y otra, a su vez, de carácter independiente. En ocasiones, según Montero (2016), esta relación es insuficiente para explicar o comprender aquellos fenómenos más complejos en lo que intervienen más de dos variables.

En la actualidad, en las evaluaciones educativas de cualquier tipo, se encuentran asociados unos cuestionarios de contexto que normalmente son rellenados por las familias de los alumnos que van a realizarlas. Su objetivo es dar a conocer el carácter socioeconómico y cultural de cada centro y de los alumnos de forma individual, entre otras funciones. Siendo así, según Gamazo et al. (2018), estos factores contextuales que son conocidos gracias a los cuestionarios de familia, pueden dividirse en dos niveles: alumnos y centros educativos. En cuanto al alumno, podemos encontrar las variables que hacen referencia a su sexo, el nivel socio-económico, condición de repetidor, el nivel educativo y profesión de sus padres, etc.

Por otro lado, y en función de los centros educativos, también se investiga acerca de aquellas variables a nivel de centro que repercuten positiva o negativamente en el rendimiento de los estudiantes, aunque, existe menos unanimidad sobre los efectos significativos que estas producen (Choi de Mendizábal y Calero, 2013). Aun así, según Perry y McConney (2010), algunas investigaciones resaltan la influencia de factores relacionados con el nivel socio-económico y cultural medio del centro y la metodología utilizada en estos en los resultados de los estudios.

Normalmente, las variables contextuales que afectan a los alumnos directamente o a las escuelas donde estudian, suelen tener un efecto bastante significativo en el rendimiento académico, es decir, en los resultados finales de las evaluaciones (Gamazo et al., 2018). Es por esto que, los alumnos estudiados no se deben de considerar independientes en función de sus puntuaciones en las pruebas de evaluación, sino que estos, como hemos mencionado con antelación, están agrupados en diversos centros que, a su vez, forman parte de un contexto situacional determinado que influye en el rendimiento académico del alumno.

Esta situación hace que los análisis cuantitativos considerados como tradicionales (modelos de regresión binaria o multivariante) se queden obsoletos para este tipo de estudios, en los que es necesario usar técnicas estadísticas que permitan el análisis a diferentes niveles y que, a su vez, consideren el anidamiento de las variables objetos de estudio (Iñiguez-Berrozpe y Marcaletti, 2018).

Como podemos observar, la realidad del sistema educativo es más compleja de lo que parece y se encuentra vinculada a una gran variedad de factores, algunos controlables y otros no tanto. Por esto, es necesario que, para poder realizar investigación educativa, se lleve a cabo un análisis que contemple un enfoque amplio y considere todas aquellas variables contextuales existentes. En los últimos años se han ido perfeccionando técnicas estadísticas que tengan en cuenta todas estas cuestiones, y se han desarrollado los modelos jerárquicos lineales (Bryk y Raudenbush, 1992).

A la hora de trabajar con muestras reales, cuyos datos no son independientes entre ellos y forman una estructura anidada o jerárquica, la mejor opción es utilizar técnicas estadísticas que permitan considerar los diferentes niveles de análisis. Concretamente, los llamados modelos jerárquicos lineales porque son capaces de ofrecernos una respuesta válida a situaciones en las que para una variable dependiente influyen más de una variable independiente (Montero, 2016).

Los modelos jerárquicos lineales aparecieron en el Reino Unido a finales de los años ochenta con el estudio *Statistical modelling issues in school effectiveness studies* de Murray Aitkin y Nicholas Longford (González, 2015). Ambos dieron al mundo de la estadística una nueva perspectiva y, actualmente, los modelos jerárquicos lineales son utilizados en investigaciones o estudios de rendimiento y eficacia escolar, ya que es necesario considerar todos los niveles de forma jerarquizada y esta es, hasta el momento, la única técnica estadística que lo proporciona.

Bryk y Raudenbush (1992) piensan que los modelos jerárquicos lineales son una válida opción para la realización de estudios de rendimiento escolar puesto que, gracias a sus características como modelo anidado, multivariante y multinivel, se puede controlar el efecto que las variables contextuales ejercen sobre el rendimiento individual del alumno y total del centro escolar. Los modelos jerárquicos lineales son técnicas estadísticas que ofrecen al usuario investigador la oportunidad de considerar la estructura anidada de los datos y respetar la jerarquía de las covariables para estudiar de forma simultánea los efectos que ejercen las variables contextuales en los diferentes niveles de análisis (alumnos y centros).

Gracias a estas técnicas, es posible eliminar la influencia que factores contextuales significativos (covariables) tienen sobre el rendimiento de estudiantes y escuelas porque se considera que una misma variable puede tener correlación o estar inducida por otras variables, por lo que existen diferentes niveles de estructuración o jerarquización (Montero, 2016). De esta forma, obtenemos un índice que informa sobre la parte del rendimiento promedio de las escuelas no explicada por estos factores contextuales. El rendimiento promedio residual obtenido en los modelos jerárquicos lineales es el que se viene denominando, de manera generalizada, como valor añadido en educación (Martínez-Abad et al., 2017).

En función del estudio, los datos pueden estar organizados en un mismo nivel (alumnos) o, en otros casos, los datos pueden tener una estructura jerarquizada o anidada en sus diferentes niveles (alumnos y centros). En otros términos, que en el análisis estadístico existan variables que hagan referencia al alumno y que, a su vez, estas, estén agrupadas en otras variables (centros educativos) con un nivel superior (Iñiguez-Berrozpe y Marcaletti, 2018).

Normalmente, en la investigación educativa basada en rendimiento académico, encontramos dos niveles característicos: alumnos y centros o escuelas. Siendo así, los alumnos se encontrarían en el nivel 1, que a su vez se encuentran anidados a un nivel superior (nivel 2) que serían los centros donde estudian. Según Iñiguez-Berrozpe y Marcaletti (2018), estos datos que nos ayudan a conocer los elementos contextuales de cada nivel y, a establecer los niveles de anidamiento necesarios, los podemos encontrar en los cuestionarios de familias que acompañan a las mismas pruebas de evaluación educativas, tal y como hemos mencionado anteriormente.

Las agencias de evaluación educativa, además de evaluar el rendimiento académico del alumnado en las diferentes competencias instrumentales, realizan preguntas para conocer las características personales, familiares y sociales del alumno. Igualmente, proporcionan datos relevantes del mismo centro como son la titularidad, método de enseñanza, formas de agrupamiento, etc. Como se puede observar, estos son datos que hacen referencia a cada uno de los niveles, por lo que son indispensables conocerlos para poder establecer conclusiones claras y realizar un estudio de rendimiento y eficacia escolar adecuado.

Los modelos jerárquicos lineales, al tener en cuenta tanto las variables asociadas a cada individuo como las contextuales que afectan al grupo, permiten comprender la varianza que produce cada nivel de agregación. De esta forma, puede analizarse qué efecto tiene la situación individual de cada alumno o alumna en su rendimiento académico y luego sumar el efecto de variables contextuales dependientes del centro, ya que los datos van introduciéndose en niveles. Así, los modelos jerárquicos lineales (HLM) permiten establecer la medida en que las correlaciones dependen del nivel individual o del nivel grupal. (Iñiguez-Berrozpe y Marcaletti, 2018, p.29).

Gracias a los modelos jerárquicos lineales se pueden llevar a cabo análisis a diferentes niveles y elaborar un modelo para cada uno de estos, en los que se consideran todas las variables que corresponden al alumnado y su agrupamiento en los distintos centros educativos. De esta forma, y, respetando la estructura anidada de los datos, comprobar de una forma más fiable la variabilidad que existen entre las covariables que se asocian a ambos niveles (González, 2015).

Por esta razón, se considera los modelos jerárquicos lineales o modelos mixtos lineales como el análisis cuantitativo más correcto para este tipo de estudios, ya que son capaces de mantener la estructura en los diferentes niveles de anidamiento para el análisis a llevar a cabo (Iñiguez-Berrozpe y Marcaletti, 2018).

Por otro lado, es importante conocer las condiciones necesarias para poder aplicar un HLM al estudio. Estas son, tal y como indica Field (2013), considerar una variable cuantitativa de tipo dependiente y que sea continua en el modelo; un número considerable de variables cuantitativas que ayuden en la predicción de los resultados; la existencia de multicolinealidad entre las variables predictoras y que, a su vez, la varianza entre ellas no sea nula; no debe darse una correlación entre las variables predictoras y las variables externas o no incluidas en el modelo multinivel. También es necesario reflexionar acerca de que los datos a analizar tendrán que seguir una estructura jerárquica o anidada. La estructura de los niveles podría ser simple, utilizando dos niveles de anidamiento o compleja, es decir, que la jerarquía de niveles comprenda tres o más niveles (Iñiguez-Berrozpe y Marcaletti, 2018).

Según Iñiguez-Berrozpe y Marcaletti (2018), es esencial que la persona o personas responsables de la investigación conozcan a la perfección la estructura de la jerarquía que van a llevar a cabo, es decir, el tipo de anidación de los datos en los diferentes niveles para que los resultados del análisis sean correctos. En este caso, nuestro anidamiento, al igual que en muchas otras investigaciones realizadas al respecto será de alumnado agrupado en centros educativos.

Por último, comentar que otra de las recomendaciones a seguir a la hora de llevar a cabo un estudio aplicando los modelos jerárquicos lineales, es el del tiempo o crecimiento. Es necesario, para proporcionar datos más fiables, analizar la duración de estos efectos de valor añadido de la eficacia escolar durante más de un año, es decir, realizar estudios longitudinales de al menos tres años para conocer el verdadero progreso y situación de eficacia de esos centros. En este caso, es aconsejado aplicar modelos longitudinales puesto que, son modelos mixtos que establecen medidas que se repiten a lo largo de los años en relación al rendimiento de los estudiantes, ayudando generar más conocimiento en el modelo (Martínez-Abad, Lizasoain, Castro y Joaristi, 2017).

3. OBJETIVOS DEL ESTUDIO

3.1. Objetivos generales

En relación a los objetivos generales que estructuran el estudio aquí presente, encontramos los siguientes:

1. Identificar y seleccionar, con la ayuda de modelos jerárquicos lineales, los centros de muy alta y muy baja eficacia escolar de la Comunidad Autónoma Andaluza considerando los criterios: promedios y crecimiento de los residuos en las competencias instrumentales básicas (razonamiento matemático y comunicación lingüística).
2. Analizar, para la elección de centros de muy alta y muy baja eficacia, los datos registrados de las evaluaciones diagnósticas contemplando una doble perspectiva basada en estudios de corte transversal contextualizado y longitudinal.

3.2. Objetivos específicos

Por otro lado, los objetivos específicos que nacen a partir de las metas inicialmente plateadas y, por lo tanto, son necesarios para lograrlas son:

1. Preparar y depurar los archivos de datos de las evaluaciones diagnósticas proporcionados por la Agencia Andaluza de Evaluación Educativa para su posterior análisis.
2. Generar la estructura jerarquizada de los datos a nivel de centros educativos y alumnos para la modelización multinivel.
3. Realizar una labor de ajuste en cuanto a los modelos jerárquicos lineales para obtener aquellas variables contextuales que influyen en el modelo.
4. Estudiar, para la selección final de centros, los resultados obtenidos atendiendo a los promedios de la diferencia residual y los residuos extremos.

4. MÉTODO

4.1. *Muestra participante*

Para la realización del estudio de selección de centros de muy alta y muy baja eficacia escolar, se han utilizado los datos derivados de las evaluaciones diagnósticas (ED) que han sido proporcionadas por la Agencia Andaluza de Evaluación Educativa (AGAEVE). Los archivos de datos corresponden a las ED que fueron cumplimentadas en los años 2013/14, 2014/15 y 2016/17, por todos los alumnos de 2º de Educación Primaria de los diferentes centros educativos de la Comunidad Autónoma Andaluza en las competencias instrumentales básicas propias del análisis: razonamiento matemático y comunicación lingüística.

Se trata entonces, de un estudio de carácter censal, transversal contextualizado y longitudinal, con una muestra muy elevada puesto que, el número de centros educativos que conforman la realidad del Sistema Educativo Andaluz es muy amplio. Esta abundante muestra de centros hace referencia a la característica estructuración del sistema educativo español y, consecuentemente, en Andalucía como dual o (trino). Un sistema educativo en el que conviven la escuela pública, la privada y la concertada. Asimismo, la muestra de centros a considerar para el estudio, debe dar respuesta a la realidad presente. Por lo tanto, no solo se analizarán los datos que provienen de los centros de titularidad pública, sino que, además, para dar validez al estudio, añadiremos los datos de los centros con titularidad privada y concertada. Siendo así, la muestra participante de centros educativos inicial ascendía a 2533.

Considerando la multitud de centros inicialmente participantes y la afluencia de datos, era necesario establecer algunos criterios para realizar un primer ajuste de la muestra que finalmente contribuirá en el estudio. Como primer criterio de ajuste, podríamos destacar el tamaño de los centros educativos, es decir, el número de alumnos que el centro tiene matriculados en el curso seleccionado (2º de Educación Primaria) y que, por tanto, han participado en las ED. En este caso, tal y como se puede apreciar en la Tabla 1, no se tuvo en cuenta, como centros relevantes para el estudio, aquellos en los que su tamaño fuera muy pequeño puesto que, las condiciones de estos están muy alejadas al resto de centros participantes y, la validez del estudio no sería precisa.

Se han considerado como centros pequeños, aquellos cuyo número de alumnos matriculados en 2º de Educación Primaria es menor a 12 estudiantes, incluyendo estos últimos. La reducción final de centros en relación a este criterio y, en función de los años analizados lo podemos apreciar en la Tabla 1.

Tabla 1.

Reducción de centros escolares participantes según su tamaño en 2º de Educación Primaria.

AÑO	Nº DE ALUMNOS	Nº DE CENTROS	DISTRIBUCIÓN
<i>Curso escolar</i> 2013/14	< 12	410	16,3 %
	≥ 12	2115	83,7 %
	TOTAL CENTROS	2525	100 %
	TOTAL PARTICIPANTE	2115	83,7%
<i>Curso escolar</i> 2014/15	< 12	395	15,6 %
	≥ 12	2138	84,4 %
	TOTAL CENTROS	2533	100 %
	TOTAL PARTICIPANTE	2138	84,4 %
<i>Curso escolar</i> 2016/17	< 12	433	17,2 %
	≥ 12	2092	82,8 %
	TOTAL CENTROS	2525	100 %
	TOTAL PARTICIPANTE	2092	82,8 %

Tal y como se puede observar en la Tabla 1, el porcentaje de centros que en los tres años evaluados tienen un número de alumnos menor que 12 en la etapa educativa señalada es inferior al 20%. Por lo que, teniendo en cuenta la amplia muestra con la que se dispone para el estudio, no causará estragos. Siendo así, se eliminaron de la muestra todos los centros que tuvieran un tamaño inferior a 12 estudiantes en 2º de primaria.

Por otra parte, un segundo criterio para el ajuste de la muestra participante, habiendo retraído el efecto del tamaño de los centros, hace referencia a las propias características del estudio longitudinal contextualizado que se está desarrollando. En el estudio aquí indicado, se analizará el crecimiento del promedio de los residuos de los centros en los tres años seleccionados (2013/14, 2014/15 y 2016/17). Siendo así, se estudiarán solo aquellos centros que hayan participado en los 3 años considerados anteriormente. Se excluirán los que solo han participado en uno o dos de los años de aplicación de la ED.

Tal y como se puede observar en la Tabla 2, la muestra participante final del estudio consta de 1936 centros escolares, considerando entre ellos las diferentes titularidades presentes: pública, privada y concertada.

Tabla 2.

Muestra final considerando los criterios de tamaño y participación en los 3 años de ED.

AÑO	CENTROS PARTICIPANTES
<i>Curso escolar 2013/14</i>	2115
<i>Curso escolar 2014/15</i>	2138
<i>Curso escolar 2016/17</i>	2092
TOTAL DE CENTROS QUE PARTICIPAN LOS 3 AÑOS	1936

Por último, comentar que los datos proporcionados por la Agencia Andaluza de Evaluación Educativa son privados y de limitada divulgación por su carácter legal. Las puntuaciones que obtuvieron cada centro educativo en las ED se encuentran bajo secreto estadístico. Siendo así, no es posible hablar abiertamente del nombre de los centros que finalmente sean seleccionados. Los centros, además, se encuentran codificados por un número de identificación que sólo la AGAEVE conoce, por lo que, los resultados del estudio serán generales y anónimos.

4.2. Instrumentos y variables

Entre los instrumentos utilizados para llevar a cabo el estudio de selección de centros, podemos encontrar dos herramientas indispensables y que han sido claves en el proceso. Estas, hacen referencia tanto a las puntuaciones como al contexto de ambos niveles de estudio: alumno y centro educativo al que pertenecen.

Por una parte, en relación a las puntuaciones, como se ha mencionado anteriormente, se han utilizado los archivos de datos que contenían los resultados de los alumnos de 2º de Educación Primaria durante los cursos 2013/14, 2014/15 y 2016/17 en las evaluaciones diagnósticas realizadas por la AGAEVE. Siendo así, se dispone de las puntuaciones en razonamiento matemático y comunicación lingüística, es decir, las dos competencias básicas analizadas.

Por otra parte, comentar que, la realidad educativa contempla una serie de factores contextuales que influyen de una forma u otra a las puntuaciones que estos alumnos obtienen y, en consecuencia, la posición que los centros educativos alcancen. Esta información de carácter social, cultural y económico las encontramos en los cuestionarios de contexto o (de familia) que se encuentran adjuntos a las pruebas de evaluación diagnóstica. Estos, son rellenados por las familias del alumnado que participa en la ED. Los cuestionarios de familia están orientados a dar a conocer el carácter socioeconómico y cultural que presenta cada centro en cuestión y, además, hacernos entender cuáles son las características del propio alumno que, inevitablemente, intervienen de forma positiva o negativa en la consecución de los resultados.

La AGAEVE facilitó de forma adjunta a los archivos de datos referentes a la ED, los cuestionarios de familia que se realizaron para cada año. Gracias a estos, se favoreció el proceso de selección final de centros de muy alta y muy baja eficacia escolar puesto que, se pudo utilizar las técnicas estadísticas de modelización multinivel y, controlar el efecto que las variables contextuales ejercen en el aprendizaje del alumno y en los resultados del centro. Entre los datos que recoge estos cuestionarios de contexto familiar podemos enumerar los siguientes: nivel socioeconómico y cultural de las familias del alumnado, año de nacimiento del alumno, sexo, el compromiso lector, la implicación de

las familias para con sus hijos en el aprendizaje, el tiempo que el alumno dedica a usar pantallas (recursos tecnológicos como ordenadores o televisión), actividades extraescolares que realiza, etc.

Sobre la información relativa a los centros educativos, no se han obtenido cuestionarios de contextos rellenos por los responsables de la institución educativa que hagan referencia al clima escolar, la propia organización del centro, modelos de aprendizaje o metodología de enseñanza utilizadas en las aulas. La única información relevante que se consiguió al respecto es la titularidad, es decir, si son de carácter público, privado o concertado y la información relativa al centro educativo presente en los cuestionarios de familia. En estos, se preguntaba a los familiares su opinión acerca del clima escolar del centro, su funcionamiento y organización.

Desafortunadamente, los datos referentes al clima y al funcionamiento del centro no serán incluidos como covariables del estudio, debido a que estas, no tienen un carácter puramente contextual. Según Joaristi et al. (2014), el clima o la metodología utilizada dependen de las políticas propias del centro y estos, pueden decidir y actuar sobre cuál es el método que llevarán a cabo o cómo es el clima de convivencia que existe en sus aulas y, en el propio centro en general. Con esto no se está diciendo que el clima o la metodología aplicada en el centro no repercute en el rendimiento final del alumnado, sino que estos factores son decisión, en cierta parte, de las estructuras y políticas de cada escuela y, por lo tanto, no tienen un componente estrictamente contextual como es, por ejemplo, el ISEC o el tiempo de uso de pantallas.

Una vez se recibió toda la información y, por tanto, todos los archivos de datos correspondientes a las evaluaciones diagnósticas y sus respectivos cuestionarios de familia en los 3 años seleccionados, tocó realizar un segundo ajuste o depuración de aquellas variables que incluiríamos en el modelo. Se tomó la decisión de incluir en los modelos jerárquicos lineales propios del análisis estadístico multinivel, solo las variables de carácter contextual, es decir, en las que el propio centro no tiene una capacidad de actuación. Todo esto, con el fin de obtener el residuo real y contrastado, correspondiente a cada una de las dos competencias evaluadas en los 3 años de análisis. Siendo así, las variables referentes a los estudiantes tendrán un componente puramente contextual y las propias covariables que caractericen a las escuelas, serán variables agregadas que

generaremos a partir del nivel 1 (estudiantes) para el nivel 2 (centros educativos). Se contará, por ejemplo, con el nivel socioeconómico y cultural (ISEC) que presenta el centro educativo, ya que este no depende del procedimiento de enseñanza o metodologías usadas en el centro, sino que es una variable que refiere únicamente al contexto en el que este se encuentra.

A continuación, se muestran todas aquellas variables que han sido consideradas, y, por consiguiente, introducidas en los modelos multinivel. A estas, las llamaremos covariables, puesto que hacen alusión al contexto situacional a nivel de alumno y centro. Asimismo, las covariables del *Nivel 1* y que se refieren a los estudiantes de los centros evaluados son las siguientes:

- ***Sexo_NI***. Género del estudiante, considerando el valor 0 para la condición de ser hombre y 1 para la condición de mujer.
- ***ISEC_NI***. Índice Socioeconómico y Cultural familiar y, por lo tanto, aplicado al estudiante. Este ha sido calculado mediante el análisis factorial de aquellos indicadores relevantes presentes en el cuestionario de familia. En primer lugar, los factores o componentes que se han tenido en cuenta para construir la variable ISEC son los que aparecen a continuación: la existencia en casa de recursos como internet, lavavajillas, programas educativos online, DVD's, el número de libros que poseen, la disposición de una habitación individual para su hijo/a y un sitio adecuado para facilitar el estudio. En segundo lugar, además de considerar los aspectos mencionados anteriormente, es necesario tener en cuenta aquellos que hacen eco de la posición de la familia según la profesión y el nivel de estudios. Tras el análisis factorial de todos los factores, se realizó una función de reemplazo de los valores perdidos por la media ya que, de esta forma, se dio la posibilidad de eliminar todos aquellos casos en los que la variable ISEC no tenía información y apareciera como 0. Siendo así, sus valores se estandarizaron considerando la media 0 y la desviación típica 1; sus valores extremos son -2.53 y 1.87.
- ***Acuesta_NI***. La hora en la que los alumnos se van a dormir los días entre semana y, por supuesto, en el periodo escolar.

- ***Cantidad_tareas_NI***. Opinión de las familias acerca de la cantidad de tareas que el alumno tiene que hacer en casa.
- ***Tiempo_tareas_NI***. Tiempo en medido en minutos que el alumno dedica en casa a la realización de las tareas o deberes escolares.
- ***Compromisolector_NI***. Esta covariable hace referencia a una suma de las diferentes preguntas que se relacionan con el concepto y, por tanto, crean el compromiso lector del alumno. Entre ellas podemos encontrar el tiempo que el alumno dedica a leer en casa y la implicación de las familias a realizar labores de lectura para con sus hijos.
- ***Implicolegio_NI***. Nivel de implicación que tienen los padres en centro escolar en el cual sus hijos se encuentran matriculados.
- ***Expectativasfam_NI***. Nivel de aspiración que tienen las familias en cuanto a la educación de sus hijos.
- ***Extraescolares_NI***. Número y tipo de actividades extraescolares a las que el alumno está inscrito.
- ***Pantallas_NI***. Tiempo que dedica el alumno en casa a ver la televisión o utilizando el ordenador o cualquier otro dispositivo tecnológico como móvil o tablet.
- ***Impliestudiante_NI***. Nivel de implicación que tiene el alumno en cuanto a su aprendizaje, el estudio y con el centro educativo en general.

Por otra parte, las covariables que forman parte del *Nivel 2* se crearon a partir de las variables de nivel 1, para conformar las tasas para cada escuela como proporciones y promedios de los resultados y puntuaciones de los estudiantes. Siendo así, las covariables dispuestas para el *Nivel 2* (centros educativos) fueron las siguientes:

- ***Sexo_N2***. Número de centros que segregan en función del género masculino o femenino.
- ***ISEC_N2***. Índice Socioeconómico y Cultural medio del alumnado del centro.
- ***Acuesta_N2***. Promedio de la hora en la que los alumnos del centro se van a dormir los días entre semana durante el periodo escolar.
- ***Cantidad_tareas_N2***. Tasa de alumnado del centro con tareas o deberes para realizar en casa.
- ***Tiempo_tareas_N2***. Promedio del tiempo que dedican los alumnos del centro a la realización de las tareas o deberes escolares.
- ***Compromisolector_N2***. Promedio del nivel de compromiso de los alumnos del centro en relación a la lectura.
- ***Implicolegio_N2***. Promedio del nivel de implicación que tienen las familias de los alumnos del centro escolar.
- ***Expectativasfam_N2***. Promedio de las expectativas de las familias de los alumnos pertenecientes a los centros educativos en relación a su formación educativa.
- ***Extraescolares_N2***. Promedio de la cantidad de actividades extraescolares a las que está adscrito el alumnado del centro educativo.

- ***Pantallas_N2***. Promedio del tiempo que dedican los alumnos del centro en casa a ver la televisión o a utilizar cualquier dispositivo tecnológico como ordenadores o móviles.
- ***Impliestudiante_N2***. Promedio del nivel de implicación que tiene el alumnado del centro en cuanto a su aprendizaje.

Una vez consideradas las variables contextuales a ambos niveles propias del estudio, se dispuso a su incorporación en los modelos jerárquicos lineales. Esta técnica estadística inherente a los modelos multiniveles, nos ayuda a controlar y retraer el efecto que las covariables indicadas ejercen sobre el rendimiento y el aprendizaje del alumnado de los centros analizados y, por tanto, a conocer el nivel de correlación que cada una de ellas presentan con los resultados finales. Mediante la realización de los HLM, obtendremos los residuos (y su promedio) de cada escuela, para cada año de análisis (2013/14, 2014/15 y 2016/17), en las dos competencias básica evaluadas (razonamiento matemático y comunicación lingüística). Por lo tanto, contaremos con la realización de 6 modelos (3 años x 2 materias instrumentales) para la identificar los centros con mayor y menor eficacia de la Comunidad Autónoma Andaluza.

4.3. Procedimiento

Como se ha presentado anteriormente, los archivos de datos pendientes de análisis en el estudio de eficacia escolar muestran una estructura jerarquizada o anidad en función de dos niveles. En este caso, contamos en primer lugar, con las puntuaciones que obtuvieron los alumnos en las dos competencias instrumentales básicas analizadas y que, haciendo referencia a los propios alumnos, constituyen el nivel 1. En segundo lugar, poseemos los datos que organizan y asocian a cada alumno con los centros educativos a los que pertenecen y que, haciendo alusión a los centros en general, constituyen el nivel 2 de anidamiento.

Se trata, por un lado, de un estudio transversal contextualizado, en el que se recogieron los datos de las ED correspondientes a los años seleccionados y, se conocen tanto las puntuaciones de los estudiantes, como las puntuaciones de los propios centros educativos en los que son agrupados los alumnos. Para llevar a cabo el estudio bajo el punto de vista de los análisis transversales de carácter contextualizado, se utilizarán modelos jerárquicos lineales, los cuales podrán dar respuesta a la variación entre ambos niveles (alumnos y centros) y considerarán todas aquellas covariables que afectan en los resultados.

Se proponen como variables dependientes del modelo a nivel 1, las puntuaciones de cada estudiante en razonamiento matemático, por un lado, y comunicación lingüística por otro. Es decir, se realizarán dos modelos por cada año de análisis que correspondan a cada una de las variables dependientes del nivel de alumnos. Una vez se proceda a la modelización, se tratará de controlar el efecto que causan las variables contextuales, en las cuales, el propio estudiante no tiene un papel de acción. Estas, se refieren a las covariables seleccionadas en el punto anterior (X_q). Asimismo, se presenta en la Tabla 3, el modelo de regresión multinivel a nivel 1, aplicable a cada una de las variables dependientes del estudio:

Tabla 3.

Ecuación de regresión lineal múltiple para el modelo multinivel a nivel de estudiantes.

ECUACIÓN MULTINIVEL A NIVEL 1 (ESTUDIANTES)	
<i>Parte estructural</i>	$Y_{ij} = \beta_{0j} + \sum_{q=1}^q \beta_{qj} X_{qij} + r_{ij}$
<i>Parte probabilística</i>	$r_{ij} \sim N(0, \sigma^2)$

Los distintos elementos presentes en las ecuaciones, significan en el modelo:

	Y_{ij}	Puntuación que es obtenida por el estudiante i de la escuela j en cada una de las dos materias instrumentales básicas evaluadas de forma individualizada, es decir, no multivariante.
	β_{0j}	Puntuación media obtenida por cada escuela en la competencia seleccionada.
	β_{qj}	Se refiere a la influencia lineal que ejerce la covariable X_q de los estudiantes.
	X_{qij}	Puntuación que tiene el estudiante i de la escuela j en la covariable X_q seleccionada.
	r_{ij}	Hace referencia el residuo que obtiene el estudiante i de la escuela j . Es decir, la diferencia entre la puntuación que obtiene el estudiante i y la puntuación que se esperaba de él en función Hace referencia el residuo que obtiene el estudiante i de la escuela j .

De esta forma, y como hipótesis inicial o modelo nulo se considerará que el nivel de actuación de las covariables del nivel 1 es similar para todos los centros educativos. En cuanto al modelo a nivel 2, que hace referencia al centro escolar, el modelo de anidación o la estructura de la ecuación de regresión múltiple que comprende a cada competencia analizada y que, en consecuencia, nos mostrará la variación entre las escuelas es, tal y como se aprecia en la Tabla 4:

Tabla 4.

Ecuación de regresión lineal múltiple para el modelo multinivel a nivel de centros educativos.

ECUACIÓN MULTINIVEL A NIVEL 2 (CENTROS)	
<i>Parte estructural</i>	$\beta_{0j} = \gamma_{00} + \sum_{s=1}^S \gamma_{0s} W_{sj} + u_{0j}$
<i>Parte probabilística</i>	$u_{0j} \sim N(0, \tau_{00})$

Los distintos elementos presentes en las ecuaciones, significan en el modelo:

{	β_{0j}	Hace referencia al rendimiento promedio de cada centro educativo j .
	γ_{00}	Indica el efecto análogo presente en todos los centros.
	γ_{0s}	Se refiere a la influencia lineal que ejerce la covariable W_s en el rendimiento medio que obtienen los centros educativos j .
	W_{sj}	Es el valor que adquiere el centro j en la covariable de la escuela W_s .
	u_{0j}	Hace referencia a la variación residual entre centros una vez hayan sido retraídos todos los efectos de las covariables consideradas en el modelo a nivel individual y de centro.

Asimismo, el modelo de carácter mixto o combinado que se llevará a cabo, y que, por ende, considera ambos niveles de análisis, corresponde con la siguiente ecuación:

$$Y_{ij} = \gamma_{00} + \sum_{s=1}^s \gamma_{0s} W_{sj} + \sum_{q=1}^q \beta_{qj} X_{qij} + u_{0j} + r_{ij}$$

La utilización, por tanto, de este modelo estadístico multinivel, se llevará a cabo con cada competencia evaluada (razonamiento matemático y comunicación lingüística) y, a su vez, para cada año de análisis (2013/14, 2014/15 y 2016/17). Concretamente, se efectuarán 6 modelos jerárquicos lineales, con los cuales, teniendo en cuanta las covariables para cada nivel de análisis, calcularemos el residuo para cada centro en cada materia de estudio. El elemento u_{0j} corresponde al residuo de la escuela j , una vez se haya controlado los efectos de las variables contextuales (covariables) que intervienen a

ambos niveles (alumnos y escuelas). Siendo así, este residuo u_{0j} , es la diferencia que encontramos entre el rendimiento medio obtenido por las escuelas y, su rendimiento esperado, una vez controlado la influencia de las covariables.

Se podría concluir con que este, es un válido indicador de eficacia escolar y un claro criterio para la selección de centros final puesto que, un centro que obtiene un promedio residual, en ambas materias evaluadas, muy alto y positivo, significa que es un centro que está desempeñando su labor educativa mejor de lo esperado. Es decir, en este caso se está indicando que es un centro de muy alta eficacia. De la misma forma, para la selección de centros de muy baja eficacia, se tendrán en cuenta aquellos cuyos residuos sean muy bajos y de valor negativo.

Dejando atrás el estudio transversal contextualizado que se ha descrito hasta el momento, y que, por un lado, nos servirá como criterio para la selección de centros en función del promedio de los residuos extremos, contamos también con un estudio de carácter longitudinal. Como un segundo criterio de selección, se planteó el hecho de que los centros de muy alta eficacia, además de considerarse aquellos que obtienen los residuos (promedios) más altos y positivos en los HLM, también lo serán aquellos cuyos residuos muestren una acusada tendencia creciente en el estudio longitudinal. De manera análoga, los centros de muy baja eficacia en el modelo longitudinal serán los que tengan una acusada tendencia decreciente en sus residuos para los años de análisis.

Para el estudio longitudinal, se utilizarán, al igual que con el estudio transversal contextualizado, los resultados de las ED en los 3 años consecutivos que la AGAEVE tiene registrados (2013/14, 2014/15 y 2016/17). Este, a su vez, solo puede llevarse a cabo en función de centros, ya que las puntuaciones de las ED son datos de cortes trasversales en las cohortes de 2º de Educación Primaria. De esta forma, mediante modelos jerárquicos multinivel, se examinarán los datos para conocer el crecimiento o decrecimiento e identificar los centros de muy alta o muy baja eficacia considerando, para el nivel 1 de anidamiento el de los tiempos (3 años) y para el nivel 2 el de los centros educativos. Las variables dependientes, en este caso, serán los residuos contextualizados y las puntuaciones que obtienen los centros educativos en las dos competencias evaluadas.

A continuación, tal y como se puede observar en la Tabla 5, las ecuaciones de regresión multinivel para los modelos longitudinales es la siguiente:

Tabla 5.

Ecuación de regresión lineal múltiple para el estudio longitudinal a nivel 1 (tiempo) y nivel 2 (centros).

	RESIDUOS	PUNTUACIONES
NIVEL 1 <i>Tiempo (años)</i>	$CR_{tj} = \pi_{0j} + \pi_{1j}AÑO_{tj} + e_{tj}$	$CP_{tj} = \pi_{0j} + \pi_{1j}AÑO_{tj} + e_t$
NIVEL 2 <i>Centros educativos</i>	$\pi_{0j} = \beta_{00} + u_{0j}$ $\pi_{1j} = \beta_{10} + u_{1j}$	$\pi_{0j} = \beta_{00} + u_{0j}$ $\pi_{1j} = \beta_{10} + u_{1j}$

De esta forma, se preparará un archivo de datos con todas las variables agregadas (nivel 2) a partir de las variables de nivel 1. Posteriormente, se llevará a cabo un proceso de estructuración de las variables a casos obteniendo, para cada centro educativo y año, una sola fila que haga referencia a las puntuaciones y otra, que muestre los residuos. En total, se tendrá 4 agrupamientos para cada centro y año, dos grupos que aluden a la competencia de razonamiento matemático (puntuaciones y residuos) y otros dos para la competencia de comunicación lingüística (puntuaciones y residuos).

Tras esto, se pasará a desarrollar el estudio de crecimiento con HLM con cada una de las variables dependientes (puntuaciones y residuos de lengua y matemáticas), mediante la sintaxis de Curran-Bauer, que nos aportarán los datos necesarios para la selección de centros final bajo el criterio de crecimiento o (decrecimiento).

4.4. Análisis de datos

Para el análisis de los datos, comenzaremos con la modelización multinivel (HLM) para realizar la labor de control y selección de las covariables que afectan en el rendimiento final, tanto del alumno como de los centros educativos. Según el estudio transversal contextualizado, se aplicaron 6 modelos jerárquicos lineales, considerando, para cada uno de ellos y, de forma independiente, los diferentes años de estudio y las dos competencias instrumentales básicas. De forma simultánea, todos comparten la estructura jerarquizada de los datos en las que el nivel 1 hace referencia a los estudiantes y el nivel 2 a los centros o escuelas a las que estos pertenecen. En cuanto a estas últimas (nivel 2), recalcar que son agregadas a partir de las del nivel 1, añadidas como covariables de efectos fijos y sin considerar los efectos aleatorios o pendientes. A continuación, se muestran los resultados de los HLM en función del año y las competencias:

Tabla 6.

Modelo multinivel para el año 2013/14 en la competencia de comunicación lingüística.

PARÁMETRO	ESTIMACIÓN	ERROR ESTÁNDAR	GL	T	SIG.	INT. DE CONFIANZA AL 95%	
						L.INFERIOR	L.SUPERIOR
Interceptación	330,596157	28,293325	2241,01	11,6	,000	275,11229	386,080020
Sexo_N1	19,836126	,597680	76623,5	33,1	,000	18,664676	21,007575
Acuesta_N1	2,290569	,429112	76073,7	5,33	,000	1,449511	3,131627
Comprolector_N1	,583568	,114825	76218,6	5,08	,000	,358511	,808625
Pantallas_N1	3,006861	,296688	76088,6	10,1	,000	2,425353	3,588368
Extraescolares_N1	4,632245	,372009	77162,5	12,4	,000	3,903110	5,361381
Expectativafam_N1	5,316830	,120685	76211,4	44,0	,000	5,080289	5,553371
ISEC_N1	19,327165	,383607	77905,9	50,3	,000	18,575298	20,079032
Implicolegio_N1	,371559	,101390	77339,4	3,66	,000	,172836	,570283
Impliestudiante_N1	-,916788	,171632	77216,1	5,34	,000	-1,253186	-,580391
Acuesta_N2	13,503457	4,986258	2174,48	2,70	,007	3,725127	23,281786
Comprolector_N2	6,511387	1,706532	2232,12	3,81	,000	3,164831	9,857942
Pantallas_N2	-11,812376	3,815200	2232,32	3,09	,002	19,294087	-4,330665
Expectativafam_N2	4,923806	1,307918	2217,059	3,76	,000	2,358934	7,488677

Tabla 7.

Modelo multinivel para el año 2013/14 en la competencia de razonamiento matemático.

PARÁMETRO	ESTIMACIÓN	ERROR ESTÁNDAR	GL	T	SIG.	INT. DE CONFIANZA AL 95%	
						L.INFERIOR	L.SUPERIOR
Interceptación	325,094414	21,055964	2296,86	15,4	,000	283,803725	366,385104
Acuesta_N1	2,483401	,427765	76752,2	5,80	,000	1,644985	3,321817
Comprolector_N1	,457262	,114847	76033,6	3,98	,000	,232163	,682362
Pantallas_N1	3,950721	,295931	76662,8	13,3	,000	3,370699	4,530743
Extraescolares_N1	4,885756	,372119	76977,9	13,1	,000	4,156404	5,615108
Expectativafam_N1	5,163389	,120704	75997,8	42,7	,000	4,926811	5,399968
ISEC_N1	18,593751	,386728	76107,3	48,0	,000	17,835765	19,351736
Implicolegio_N1	,475913	,101441	77070,5	4,69	,000	,277089	,674737
Impliestudiante_N1	-,878074	,171711	77038,0	5,11	,000	-1,214627	-,541521
Sexo_N1	1,678944	,597639	76488,8	2,80	,005	,507574	2,850313
Comprolector_N2	6,608611	1,797599	2224,04	3,67	,007	3,083463	10,133759
Expectativafam_N2	6,104488	1,644383	2197,23	3,71	,000	2,879779	9,329197
ISEC_N2	-7,631103	2,695832	2151,21	-2,8	,005	-12,917811	-2,344394
Tamaño_N2	-,132687	,051085	1983,42	-2,5	,009	-,232873	-,032501

Tabla 8.

Modelo multinivel para el año 2014/15 en la competencia de comunicación lingüística.

PARÁMETRO	ESTIMACIÓN	ERROR ESTÁNDAR	GL	T	SIG.	INT. DE CONFIANZA AL 95%	
						L.INFERIOR	L.SUPERIOR
Interceptación	345,175746	23,172930	2233,6	14,8	,000	299,733015	390,618478
Sexo_N1	27,342167	,585883	78685,9	46,6	,000	26,193839	28,490494
Acuesta_N1	3,135709	,411604	79037,9	7,6	,000	2,328968	3,942450
Cantidadtarear_N1	-3,037596	,539559	77983,8	-5,6	,000	-4,095129	-1,980063
Tiempotareas_N1	-16,849459	,370608	78012,2	-45,4	,000	-17,575847	-16,123071
Comprolector_N1	,868831	,113373	78729,6	7,6	,000	,646619	1,091042
ISEC_N1	18,214913	,377083	79910,0	48,3	,000	17,475833	18,953993
Implicolegio_N1	,334273	,092982	78040,0	3,5	,000	,152029	,516516
Expectfam_N1	5,378503	,117681	78140,8	45,7	,000	5,147850	5,609156
Extraescolares_N1	4,663926	,353509	79203,9	13,1	,000	3,971051	5,356801
Pantallas_N1	2,146659	,288974	78058,9	7,4	,000	1,580272	2,713047
Cantidadtarear_N2	17,707167	5,661694	2225,7	3,1	,002	6,604413	28,809921
Tiempotareas_N2	22,046879	2,961170	2245,4	7,4	,000	16,239962	27,853795
Implicolegio_N2	2,377139	,802383	2229,7	2,9	,003	,803643	3,950636
Expectfam_N2	4,353785	1,100412	2309,0	3,9	,000	2,195886	6,511684
Pantallas_N2	-11,828561	3,377124	2213,5	-3,5	,000	-18,451223	-5,205899

Tabla 9.

Modelo multinivel para el año 2014/15 en la competencia de razonamiento matemático.

PARÁMETRO	ESTIMACIÓN	ERROR ESTÁNDAR	GL	T	SIG.	INT. DE CONFIANZA AL 95%	
						L.INFERIOR	L.SUPERIOR
Interceptación	373,830290	23,801158	2209,3	15,7	,000	327,155306	420,505273
Sexo_N1	3,556726	,586750	78532,5	6,0	,000	2,406700	4,706753
Acuesta_N1	2,516247	,412392	78827,6	6,1	,000	1,707961	3,324534
Cantidadtareas_N1	-3,168897	,539991	77946,3	-5,8	,000	-4,227276	-2,110518
Tiempotareas_N1	-14,111769	,371046	77969,9	-38,0	,000	-14,839017	-13,384522
Comprolector_N1	,737988	,113524	78568,2	6,5	,000	,515481	,960496
ISEC_N1	15,851671	,378233	79627,4	41,9	,000	15,110336	16,593006
Implicolegio_N1	,331662	,092635	79225,3	3,5	,000	,150099	,513225
Expecfam_N1	5,048059	,117824	78084,7	42,8	,000	4,817125	5,278992
Extraescolares_N1	5,336367	,354090	78980,0	15,0	,000	4,642353	6,030380
Pantallas_N1	2,477887	,289211	78008,0	8,5	,000	1,911035	3,044739
Cantidadtareas_N2	22,935659	6,302248	2204,6	3,6	,000	10,576694	35,294624
Tiempotareas_N2	23,658267	3,268191	2227,5	7,2	,000	17,249248	30,067286
Expecfam_N2	3,252573	1,238554	2263,5	2,6	,009	,823753	5,681394
Pantallas_N2	-11,530242	3,713339	2194,5	-3,1	,002	-18,812269	-4,248216
Tamaño_N2	-,140082	,050900	2011,3	-2,7	,006	-,239905	-,040260

Tabla 10.

Modelo multinivel para el año 2016/17 en la competencia de comunicación lingüística.

PARÁMETRO	ESTIMACIÓN	ERROR ESTÁNDAR	GL	T	SIG.	INT. DE CONFIANZA AL 95%	
						L.INFERIOR	L.SUPERIOR
Interceptación	254,513180	23,730500	2170,1	10,7	,000	207,976299	301,050061
Sexo_N1	28,281017	,590041	73782,3	47,9	,000	27,124538	29,437495
Acuesta_N1	2,232274	,415687	73801,2	5,3	,000	1,417529	3,047019
Cantidadtareas_N1	-2,550046	,480498	73770,7	-5,3	,000	-3,491820	-1,608271
Tiempotareas_N1	-17,436674	,361797	73831,5	-48,1	,000	-18,145796	-16,727553
Comprlector_N1	,699937	,115699	73826,6	6,0	,000	,473167	,926707
ISEC_N1	14,787308	,384938	73878,2	38,4	,000	14,032831	15,541786
Extraescolares_N1	5,086626	,360160	74935,9	14,1	,000	4,380714	5,792538
Pantallas_N1	,937706	,250777	74487,2	3,7	,000	,446185	1,429227
Expecfam_N1	4,744617	,120911	74543,1	39,2	,000	4,507632	4,981603
Impliestudiante_N1	-,443114	,155033	74750,7	-2,8	,004	-,746977	-,139251
Implicolegio_N1	,605707	,097839	73820,8	6,1	,000	,413944	,797470
Sexo_N2	-24,997808	8,851277	2114,7	-2,8	,005	-42,355926	-7,639690
Acuesta_N2	14,282533	4,756508	2113,7	3,0	,003	4,954607	23,610459
Cantidadtareas_N2	23,902301	5,924613	2156,0	4,0	,000	12,283751	35,520851

Tiempotareas_N2	21,139645	3,322286	2198,7	6,3	,000	14,624497	27,654793
comprlector_N2	5,051677	1,627810	2171,1	3,1	,002	1,859449	8,243906
ISEC_N2	10,961740	2,493784	2167,8	4,3	,000	6,071282	15,852198
implicolegio_N2	2,731217	,773370	2174,9	3,5	,000	1,214596	4,247839

Tabla 11.

Modelo multinivel para el año 2016/17 en la competencia de razonamiento matemático.

PARÁMETRO	ESTIMACIÓN	ERROR ESTÁNDAR	GL	T	SIG.	INT. DE CONFIANZA AL 95%	
						L.INFERIOR	L.SUPERIOR
Interceptación	279,860929	23,441388	2206,7	11,9	,000	233,891440	325,830418
Sexo_N1	3,034363	,598280	74360,9	5,0	,000	1,861737	4,206988
Acuesta_N1	2,878425	,422436	73831,4	6,8	,000	2,050453	3,706396
Cantidadtarear_N1	-2,566358	,488182	73779,7	-5,2	,000	-3,523193	-1,609522
Tiempotareas_N1	-16,506893	,367572	73882,6	-44,9	,000	-17,227333	-15,786453
Comprlector_N1	,518292	,117484	74044,5	4,4	,000	,288023	,748561
ISEC_N1	15,330662	,387464	75677,7	39,5	,000	14,571236	16,090089
Extraescolares_N1	5,348338	,365976	74919,7	14,6	,000	4,631027	6,065649
Pantallas_N1	1,622097	,254737	74534,3	6,3	,000	1,122814	2,121380
Expecfam_N1	4,712849	,122819	74564,0	38,3	,000	4,472124	4,953575
Implicolegio_N1	-,971935	,157570	74742,9	-6,1	,000	-1,280772	-,663098
Implicolegio_N1	,775588	,099379	73958,6	7,8	,000	,580806	,970370
Acuesta_N2	12,111705	4,618977	2135,8	2,6	,009	3,053543	21,169867
Cantidadtarear_N2	26,414861	5,819747	2180,3	4,5	,000	15,002032	37,827691
Tiempotareas_N2	16,872340	3,211456	2205,1	5,2	,000	10,574544	23,170135
Comprlector_N2	5,019072	1,469704	2224,1	3,4	,001	2,136937	7,901207
implicolegio_N2	2,110272	,785667	2190,8	2,6	,007	,569542	3,651002

Para todos los modelos, tal y como se pueden apreciar en las Tablas anteriormente presentadas, solo se han considerado aquellas covariables que resultaban relevantes o verdaderamente significativas para cada modelo, es decir, que resolvían y explicaban la influencia de los factores contextuales en el rendimiento. Fueron eliminadas todas las covariables referentes a cada nivel que no mostraran un valor de significación importante. Se consideró que, el valor de significación (p) tendría que ser menor que 0,01, para así, poder afirmar con un 99% de confianza que las correlaciones o la influencia de las variables contextuales señaladas son de gran envergadura a la hora de explicar el rendimiento de los centros.

Por otro lado, y en cuanto a la interpretación de los modelos, es importante recalcar que están planteados teniendo en cuenta que todas las covariables se miden desde el nivel 0, y desde este punto, la explicación resulta con la primera columna que muestra las estimaciones. Haciendo referencia, por ejemplo, a la Tabla 11, el modelo indica que la media en razonamiento matemático para los chicos en el curso 2016/17 es de 279,86 (intercepción), en la cual, todas las demás covariables se consideran nulas y, por cada punto de incremento en cualquiera de las variables, se suma la cantidad indicada a la puntuación. En este caso, el hecho de ser chica (columna de estimación en covariable `sexo_N1`), hace que se sume la cantidad de 3,03 a la puntuación final. En las covariables cuya estimación es negativa, ocurre lo mismo, pero en este caso, por cada punto de incremento se resta la cantidad señalada. Un ejemplo lo tenemos en la Tabla 11, en el que para el modelo de matemáticas del curso 2016/17, por cada punto de incremento en la covariable referente a la cantidad de tareas, hace que la puntuación final sea inferior en (2, 56).

Una vez realizados los modelos y, habiendo controlado los efectos que las variables contextuales ejercen, se pasó al cálculo de los residuos para cada estudiante en cada una de las competencias y para cada año de análisis. Este se obtuvo a partir de la diferencia entre las puntuaciones medias previstas y las obtenidas según el modelo jerárquico lineal. Posteriormente, al generar el archivo agregado para la selección final de los centros en función del criterio del promedio de los residuos extremos, se crearon los residuos correspondientes a cada centro educativo. Además, se valoró la decisión de ordenar los residuos en función de centiles a través de una agrupación visual. De esta forma, para el residuo de cada competencia evaluada, los puntos de corte se organizarán y será más fácil la selección. Sin embargo, es importante mencionar que no se establece ningún punto de corte específico en cuanto a centiles, ya que la selección iba a constar de 25 CAEF y 25 CBEF según el criterio de los residuos extremos, independientemente de los centiles provistos. Siendo así, se obtuvieron dos puntuaciones determinantes: residuos y centil de los residuos.

Por último, para terminar con el estudio transversal contextualizado y teniendo el residuo de los centros para cada competencia evaluada en los diferentes años considerados, se pasó al cálculo del promedio del residuo total para cada centro en ambas competencias. Esta variable, que hace alusión al promedio de los residuos para ambas

competencias, sería decisiva para la selección final de los centros de alta o baja eficacia escolar. De la misma forma, se calculó el promedio de los residuos en centiles. Para comprobar la validez de la decisión de considerar aquellos centros eficaces en función del promedio de los residuos en centiles se pasó a correlacionar ambas variables.

Tabla 12.

Correlación de Pearson para las variables “RSDMEDIOCENTIL” y “RSDMEDIO”.

		RSDMEDIOCENTIL	RSDMEDIO
RSDMEDIOCENTIL	Correlación de Pearson	1	,980**
	Sig. (bilateral)		,000
	N	1936	1936
RSDMEDIO	Correlación de Pearson	,980**	1
	Sig. (bilateral)	,000	
	N	1936	1936

Como podemos observar en la Tabla 12, el grado de significación obtenido para el coeficiente R_{xy} (,000) es inferior a 0,01 por lo que podemos afirmar, con un nivel de confianza de 99%, que existe correlación entre las variables analizadas. Atendiendo al valor y al signo del coeficiente (,980), podemos afirmar, tomando como referencia la Tabla de valores de Bisquerra, que existe una tendencia muy alta a que ambas variables presenten la misma información. Con esto, podemos decir que la decisión de establecer un orden en los centros en función del promedio de su residuo en centiles, es válida. Siendo así, y para realizar la identificación y selección de centros que se presentará en el apartado de resultados, ordenaremos los casos considerando como variable clave “RSDMEDIOCENTIL”, ya sea en orden descendiente (para la selección de centros de muy alta eficacia) o en orden ascendente (para la selección de centros de muy baja eficacia).

Para el estudio longitudinal, se barajó el criterio de que un centro sería considerado de alta eficacia si sus residuos muestran una acusada tendencia creciente en los 3 años analizados. Análogamente, un centro de baja eficacia será aquel que sus residuos presentes una acusada tendencia decreciente en los modelos jerárquicos lineales. Siendo así, y para la obtención de estos residuos, se llevarán a cabo dos modelos estadístico multinivel, es decir, uno por cada competencia evaluada. En ambos modelos, tendremos

como nivel 1 el de los tiempos, es decir, los años de análisis (2013/14, 2014/15 y 2016/17). En cuanto al nivel 2, se considerarán los centros educativos.

Se introducirá en el modelo el factor año a efectos fijos a través de un análisis factorial del mismo y esta vez, incluiremos también el año sin estructura prevista en la parte aleatoria del modelo. Una vez tengamos los resultados de cada modelo, se llevará a cabo un proceso minucioso para el cálculo de los residuos. Este es, a través de la sintaxis de Curran-Bauer para pendientes aleatorias. De esta forma, se irán sustituyendo los datos necesarios para crear un nuevo archivo de datos que contenga solamente la identificación para cada centro con su intersección y pendiente.

Para llevar a cabo el proceso, fueron necesarios los valores referentes a la estructura de covarianzas de efectos aleatorios y las estimaciones de parámetros de covarianza para cada uno de los modelos realizados (modelo del residuo de razonamiento matemático y modelo del residuo de comunicación lingüística). En relación a las estimaciones de parámetros de covarianzas, utilizamos el valor del residuo en el modelo de comunicación lingüística que equivale a 773,660419. De la misma forma, el valor correspondiente a el residuo en el modelo de razonamiento matemático es igual a 985,528977. A continuación, se presentan las Tablas que hacen alusión a los valores restantes que se utilizaron para el estudio de crecimiento con HLM en el modelo de Curran-Bauer:

Tabla 13.

Estructura de covarianzas de efecto aleatorio para el modelo del residuo en comunicación lingüística.

	INTERCEPTACIÓN/ CENTRO	AÑO/ CENTRO
INTERCEPTACIÓN/ CENTRO	772,545090	-154,45742
AÑO/ CENTRO	-154,45742	122,353840

Tabla 14.

Estructura de covarianzas de efecto aleatorio para el modelo del residuo en razonamiento matemático.

	INTERCEPTACIÓN/ CENTRO	AÑO/ CENTRO
INTERCEPTACIÓN/ CENTRO	759,060767	-135,38807
AÑO/ CENTRO	-135,38807	87,244504

Finalmente, se fusionaron los archivos que contemplaban los residuos por competencia y se creó un archivo conjunto. El objetivo era calcular el promedio residual de ambas competencias para, posteriormente, en función de los rangos promedios, ordenar los casos y obtener la selección final de 25 CAEF y 25 CBEF, en función del criterio de crecimiento o decrecimiento del promedio de los residuos de los centros.

En última instancia, me gustaría hacer mención al porqué de la elección de los criterios de selección de los centros. En un primer momento, se pensó en contabilizar los centros en función del promedio de los residuos y puntuaciones tanto extremas, para el estudio transversal contextualizado, como el promedio de estos en función de su crecimiento para el estudio longitudinal.

Siguiendo estos criterios de selección iniciales, se encontró la problemática de contar con una selección de centros demasiado amplia y se decidió pasar a un ajuste. Este ajuste se basó en un procedimiento de correlacionar las puntuaciones con los residuos. Es decir, bajo la perspectiva del estudio transversal contextualizado, se observó la correlación entre las dos variables que inicialmente eran consideradas como esenciales para la selección de centros: promedios de las puntuaciones extremas y el promedio de los residuos extremos. Al comprobar la correlación entre ambas, los resultados mostraban que prácticamente las variables presentaban los mismos datos. A continuación, se adjunta la Tabla 15, que muestra la correlación entre las variables comentadas con anterioridad.

Tabla 15.

Correlación de Pearson para las variables “RSDMEDIO” y “PUNTMEDIO”.

		RSDMEDIO	PUNTMEDIO
RSDMEDIO	Correlación de Pearson	1	,879**
	Sig. (bilateral)		,000
	N	1936	1936
PUNTMEDIO	Correlación de Pearson	,879**	1
	Sig. (bilateral)	,000	
	N	1936	1936

Como podemos observar en la Tabla 15, el grado de significación obtenido para el coeficiente R_{xy} (,000) es inferior a 0,01 por lo que podemos afirmar, con un nivel de confianza de 99%, que existe correlación entre las variables analizadas. Atendiendo al valor y al signo del coeficiente (,980), podemos afirmar, tomando como referencia la Tabla de valores de Bisquerra, que existe una tendencia alta a que ambas variables presenten la misma información. Siendo así, se decidió considerar únicamente los residuos, ya que estos, son un criterio más fiable y con mayor valor para el estudio. De forma análoga, para el modelo longitudinal se realizaron los mismos pasos para comprobar la correlación entre el promedio de los residuos y las puntuaciones. Al pasar el coeficiente de correlación de Pearson entre ambas, tal y como se puede apreciar en la Tabla 16, se afirma con un nivel de confianza del 99% que existe una relación muy alta entre ambas variables y, por lo tanto, con prácticamente iguales. Asimismo, y al igual que en el criterio anterior, se concluyó con la selección de centros a partir del crecimiento (o decrecimiento) del promedio de los residuos, y no de las puntuaciones.

Tabla 15.

Correlación de Pearson para las variables “RANGOMEDIORS” y “PANGOMEDIOPUNT”.

		RANGOMEDIORS	RANGOMEDIOPUNT
RANGOMEDIORS	Correlación de Pearson	1	,943**
	Sig. (bilateral)		,000
	N	1936	1936
RANGOMEDIOPUNT	Correlación de Pearson	,943**	1
	Sig. (bilateral)	,000	
	N	1936	1936

5. RESULTADOS

Como se conoce, la investigación de eficacia y rendimiento escolar va, en su mayoría de veces, asociada con el estudio de la mejora y la calidad de los sistemas educativos. Es importante recalcar que, a pesar de la posible utilización de este estudio para la elaboración de un registro de prácticas eficaces que ayuden a focalizar y enriquecer el rendimiento de los estudiantes y los centros educativos, no es el objetivo principal del presente. Para llevar a cabo una investigación más certera acerca de los puntos o prácticas que conllevan a una mejora de los centros y, a potenciar las reformas educativas, es necesario ampliar el estudio con un análisis a nivel más cualitativo como podría ser el uso de entrevistas, observaciones o estudios de casos múltiples (Martínez-Arias, 2009). Desafortunadamente, por tiempo y otras limitaciones encontradas que se comentarán más adelante, no se ha podido realizar esta labor, por lo que el ámbito, así visto, sobre la mejora de la calidad educativa, se sale del margen de este trabajo de investigación.

Centrándonos en los objetivos planteados en este estudio, se quería conocer y seleccionar las escuelas de la Comunidad Autónoma Andaluza que obtienen unos resultados significativamente mejores o peores de lo que se espera, considerando su contexto intra-escuela. Para llegar a ellas, se tuvo que efectuar una labor muy importante: controlar los factores contextuales o covariables que influyen en los resultados y que, condicionan, de una forma u otra, esa puntuación “prevista” o “esperada” de la que se hablaba anteriormente. Mediante técnicas estadística multinivel, concretamente, modelos jerárquicos lineales, se ha dado respuesta al estudio transversal y longitudinal contextualizado que se planteaba ya que, en función de los niveles de análisis para cada uno de estos, se han logrado retraer el efecto de las variables contextuales y conocer así, los promedios de los residuos y residuos extremos de los centros educativos analizados.

Por lo tanto, como resultados finales a este trabajo, se mostrarán, por un lado, las covariables que han sido consideradas como predominantes en los modelos para cada competencia evaluada en los años de estudio y, por supuesto, la selección final de los CAEF y CBEF en función de los criterios: promedio de los residuos extremos y crecimiento (o decrecimiento) del promedio de los residuos.

5.1. Covariables significativas en el estudio

Como se ha mencionado anteriormente, a pesar de no haber realizado como tal, una caracterización de los centros, si se ha llegado a conocer cuáles son las variables de carácter contextual que afectan o influyen positiva o negativamente, tanto a nivel de alumnos como de centros, en el rendimiento final. A continuación, se presenta en la Tabla 16, el conjunto de covariables significativas para el modelo multinivel en los diferentes años de análisis para la competencia de comunicación lingüística a ambos niveles.

Tabla 16.

Covariables significativas en 2º de Educación Primaria para los dos niveles en la competencia comunicación lingüística.

MODELO JERÁRQUICO PARA COMUNICACIÓN LINGÜÍSTICA		
	NIVEL 1 (estudiantes)	NIVEL 2 (centros)
AÑO 2013/14	→ <i>Sexo</i>	→ <i>Promedio de la hora de acostarse</i>
	→ <i>ISEC</i>	→ <i>Promedio del nivel de compromiso lector</i>
	→ <i>Hora de acostarse</i>	→ <i>Promedio del tiempo de pantallas</i>
	→ <i>Compromiso lector</i>	→ <i>Promedio de las expectativas familiares</i>
	→ <i>Tiempo de pantallas</i>	
	→ <i>Actividades extraescolares</i>	
	→ <i>Expectativas familiares</i>	
	→ <i>Implicación con el colegio</i>	
AÑO 2014/15	→ <i>Implicación del estudiante</i>	
	→ <i>Sexo</i>	→ <i>Promedio del nivel de implicación con el colegio</i>
	→ <i>ISEC</i>	→ <i>Promedio del tiempo de pantallas</i>
	→ <i>Hora de acostarse</i>	→ <i>Promedio de las expectativas familiares</i>
	→ <i>Compromiso lector</i>	→ <i>Tasa de alumnado del centro con tareas</i>
	→ <i>Tiempo de pantallas</i>	→ <i>Promedio del tiempo que dedican a las tareas</i>
	→ <i>Actividades extraescolares</i>	
	→ <i>Expectativas familiares</i>	
	→ <i>Implicación con el colegio</i>	
	→ <i>Cantidad de tareas</i>	
→ <i>Tiempo de tareas</i>		

AÑO 2016/17

- | | |
|-------------------------------------|---|
| → <i>Sexo</i> | → <i>Número de centros que segregan en función del sexo</i> |
| → <i>ISEC</i> | → <i>Promedio de la hora de acostarse</i> |
| → <i>Hora de acostarse</i> | → <i>Tasa de alumnado del centro con tareas</i> |
| → <i>Compromiso lector</i> | → <i>Promedio del tiempo que dedican a las tareas</i> |
| → <i>Tiempo de pantallas</i> | → <i>ISEC medio del centro</i> |
| → <i>Actividades extraescolares</i> | → <i>Promedio del nivel de implicación con el colegio</i> |
| → <i>Expectativas familiares</i> | |
| → <i>Implicación con el colegio</i> | |
| → <i>Implicación del estudiante</i> | |
| → <i>Cantidad de tareas</i> | |
| → <i>Tiempo de tareas</i> | |

De la misma forma, y tal y como se puede apreciar en la Tabla 17, se presentan las covariables influyentes para el modelo en la competencia de razonamiento matemático.

Tabla 17.

Covariables significativas en 2º de Educación Primaria para los dos niveles en la competencia razonamiento matemático.

MODELO JERÁRQUICO PARA RAZONAMIENTO MATEMÁTICO	
NIVEL 1 (estudiantes)	NIVEL 2 (centros)
→ <i>Sexo</i>	→ <i>ISEC medio del centro</i>
→ <i>ISEC</i>	→ <i>Promedio del nivel de compromiso lector</i>
→ <i>Hora de acostarse</i>	→ <i>Tasa de alumnos (tamaño) en 2º de Educación Primaria.</i>
→ <i>Compromiso lector</i>	→ <i>Promedio de las expectativas familiares</i>
→ <i>Tiempo de pantallas</i>	
→ <i>Actividades extraescolares</i>	
→ <i>Expectativas familiares</i>	
→ <i>Implicación con el colegio</i>	
→ <i>Implicación del estudiante</i>	

AÑO 2013/14

AÑO 2014/15

- | | |
|-------------------------------------|--|
| → <i>Sexo</i> | → <i>Tasa de alumnos (tamaño) en 2º de Educación Primaria.</i> |
| → <i>ISEC</i> | → <i>Promedio del tiempo de pantallas</i> |
| → <i>Hora de acostarse</i> | → <i>Promedio de las expectativas familiares</i> |
| → <i>Compromiso lector</i> | → <i>Tasa de alumnado del centro con tareas</i> |
| → <i>Tiempo de pantallas</i> | → <i>Promedio del tiempo que dedican a las tareas</i> |
| → <i>Actividades extraescolares</i> | |
| → <i>Expectativas familiares</i> | |
| → <i>Implicación con el colegio</i> | |
| → <i>Cantidad de tareas</i> | |
| → <i>Tiempo de tareas</i> | |

AÑO 2016/17

- | | |
|-------------------------------------|---|
| → <i>Sexo</i> | → <i>Promedio del nivel de compromiso lector</i> |
| → <i>ISEC</i> | → <i>Promedio de la hora de acostarse</i> |
| → <i>Hora de acostarse</i> | → <i>Tasa de alumnado del centro con tareas</i> |
| → <i>Compromiso lector</i> | → <i>Promedio del tiempo que dedican a las tareas</i> |
| → <i>Tiempo de pantallas</i> | → <i>Promedio del nivel de implicación con el colegio</i> |
| → <i>Actividades extraescolares</i> | |
| → <i>Expectativas familiares</i> | |
| → <i>Implicación con el colegio</i> | |
| → <i>Implicación del estudiante</i> | |
| → <i>Cantidad de tareas</i> | |
| → <i>Tiempo de tareas</i> | |

Lo verdaderamente importante aquí, no son los modelos o su cuantía en cuanto a estimaciones o parámetros, sino ofrecer una lista, de aquellas variables que hacen referencia al contexto del centro educativo o del alumno, y afectan positiva o negativamente en los resultados que se obtengan. Como podemos observar en las Tablas 16 y 17, parece que el producto de las covariables que intervienen en los modelos, están muy relacionados con los resultados de otros estudios sobre eficacia y rendimiento escolar. Entre ellos, destacar los estudios de Santiago, Lukas, Joaristi, Lizasoain y Moyano (2008), o Joaristi et al. (2014), donde se muestra la gran influencia, a nivel 1, del género del alumno o el Índice Socioeconómico y Cultural de las familias en los resultados de los modelos de las competencias evaluadas.

Por otro lado, encontramos también diferencias importantes que hacen referencia a la realidad del Sistema Educativo Andaluz en el que, covariables como las expectativas que tienen las familias sobre sus hijos, el tiempo que dedican los alumnos al uso de pantallas (ordenadores o móviles) y la hora en la que se acuestan repercuten también en el rendimiento y la eficacia final que los centros obtienen. En los estudios anteriormente mencionados, no muestran una relación en la influencia de estos factores para el nivel de estudiantes, pudiendo decir que, el hecho de que en este estudio si exista una correlación importante, abre una nueva perspectiva para posibles investigaciones en el contexto andaluz.

En cuanto al nivel 2, que hacía referencia a las covariables que inciden en los resultados de los centros educativos, encontramos que solo en el modelo para razonamiento matemático en el año 2013/14 y para el de comunicación lingüística en 2016/17, el ISEC medio del centro es un factor relevante, mientras que, en otros estudios al respecto llevados a cabo en otros territorios comentan que, este factor es una buena variable explicativa puesto que, se encuentra presente en todos los modelos. En cambio, encontramos que covariables como la tasa de alumnado con tareas y el promedio del tiempo que dedican a su realización, si son muy relevantes. En conclusión, se puede decir que las variables contextuales que afectan a los resultados del centro a nivel 2, se encuentran más diferenciadas y, sería necesaria una labor de análisis más cualitativo para detectar verdaderamente su influencia y su importancia en la práctica educativa.

Una vez se obtuvieron, como es en este caso, según el modelo transversal contextualizado y longitudinal, todas las covariables que intervienen en los resultados para los 3 años seleccionados y que, por ende, nos ayudan a conformar ese valor añadido para comprobar la diferencia residual entre puntuaciones obtenidas y esperadas, se pasa al objetivo principal del estudio: la selección de centros.

5.2. Selección de centros

Para el objetivo principal del estudio, es decir, la selección de centros en la Comunidad Autónoma Andaluza de muy alta y muy baja eficacia, se tuvieron en cuenta dos criterios claves: el promedio de los residuos extremos calculados a partir de HLM en el estudio transversal contextualizado y los promedios de los residuos en función del crecimiento o decrecimiento calculados también por HLM en el estudio longitudinal.

Siendo así, y una vez fueron diseñados, aplicados y comprobados los modelos para el estudio tanto transversal como longitudinal, se calcularon los residuos para cada centro educativo. Este representa la diferencia controlada entre el rendimiento medio obtenido y el esperado tras realizar la labor de control o retraining de las covariables que podían influir en los modelos (Joaristi et al., 2014).

En relación al criterio de residuos extremos, se encontró la dificultad referente al tiempo (años). Los centros de muy alta eficacia, siguiendo los indicadores de este principio, serían los que obtuvieran los mejores y los peores residuos. Sin embargo, se contaba con la problemática de que, en ocasiones, los centros que en un año de estudio y competencia evaluada habían obtenido un residuo muy alto y de valor positivo, en la otra competencia o para el siguiente año, los resultados eran totalmente contrarios o, en su defecto, diferentes.

Siendo así, y considerando las dificultades que se presentan en los estudios de valor añadido realizados, se pensó en considerar centros de alta eficacia, aquellos que obtuvieran los residuos más altos y positivos en las dos materias instrumentales básicas, y no de forma independiente, ya que el efecto de los años podríamos contrastarlos con el estudio longitudinal. De esta forma, considerando aún el criterio de los residuos extremos para ambas competencias de forma conjunta, se reordenaron estos mismos residuos en función de centiles y, en cuanto a esta última variable, se seleccionaron los centros más y menos eficaces.

Por otro lado, en relación al criterio de crecimiento del promedio de los residuos mediante modelos jerárquicos lineales en el estudio longitudinal, se tuvieron en cuenta

como centros de alta eficacia: aquellos cuyos residuos, para ambas competencias evaluadas, crecieran más en los 3 años de análisis. De la misma forma, los centros de baja eficacia, serían aquellos en los que sus residuos mostraran una fuerte tendencia decreciente en los años de estudio y, para las dos competencias instrumentales evaluadas de forma conjunta. Además, y para llevar a cabo la selección, se reordenaron también, en función de rangos, los residuos calculados para cada competencia y año. El objetivo era crear una variable que reflejara el promedio de los rangos para ambas competencias, la cual se utilizó para ordenar los centros. Este fue únicamente un método de orden para la selección de centros, ya que, en ningún momento, se pensó en indicar un punto de corte en función de centiles o rangos.

La selección de centros, se llevó a cabo en función de los criterios: promedio de los residuos según crecimiento o (decrecimiento) y promedio de los residuos extremos. Siendo así, y contemplando la utilización de este estudio para investigaciones futuras, se observó la necesidad de seleccionar un número considerable de CAEF y CBEF. Normalmente, cuando se quiere realizar un estudio de tipo más cualitativo dónde es necesario la colaboración de los integrantes de los centros seleccionados, muchos de ellos, se muestran reacios o no quieren participar. Siendo así, se optó por hacer una selección de centros más o menos amplia para que si en un futuro, se llegara a dar esta situación, poder solventarla y tener más centros con los que contar en el estudio.

Finalmente, se seleccionaron 50 CAEF y 50 CBEF, de los cuales 25 de los centros considerados como de muy alta eficacia se eligieron siguiendo el criterio que contemplaba la perspectiva de los estudios transversales contextualizados: promedio de los residuos extremos. De forma análoga, se identificaron 25 centros de baja eficacia siguiendo esta misma lógica, pero en orden inversa. Por otro lado, y en relación a los 25 centros restantes para cada pauta de selección, se escogieron 25 CAEF bajo el punto de vista del estudio longitudinal, dónde el criterio aquí fue el crecimiento del promedio de los residuos en las competencias analizadas en los 3 años de estudio. Asimismo, se seleccionaron otros 25 CBEF teniendo en cuenta el decrecimiento del promedio de sus residuos. En la Tabla 18, se puede observar la selección final de centros.

Tabla 18.

Selección de centros de muy alta y muy baja eficacia considerando como criterios: promedio de los residuos según crecimiento o (decrecimiento) y residuos extremos.

CAEF RSDs EXTREMOS	CAEF CREC RSDs	CBEF RSDs EXTREMOS	CBEF CREC RSDs
1523182	5822217	8642672	7002436
5516762	3677586	5617804	5730670
647914	3760652	6463606	1948989
4245180	<u>5698449</u>	5546602	1611514
6223517	1263692	9270809	5586809
668611	7143126	8778015	2590468
4003750	5459537	<u>5698449</u>	6746715
7706836	9405522	550661	9788762
2317249	3622189	6558005	2139949
3160903	9587320	7549315	9591851
1203224	2708205	4138835	9977102
714743	6989581	510074	5199614
876706	1312788	3567468	3081393
8613584	4620220	1961836	2345458
1343211	4544166	9419455	9433949
4596269	4900223	8829862	7181757
6648673	5896380	2160114	4371335
6022678	7497009	4465472	5290892
3843757	1414378	3666265	7199641
8177535	5812243	1968217	6239322
8850997	6811263	9302288	5897527
8755573	2877200	9703769	2424556
8958894	4150316	2114249	8417745
4082480	7770949	4671231	3249667
5251619	5468663	7332829	5693555

Nota: Las cifras aquí presentadas hacen alusión a los identificadores (proporcionadas por la AGAEVE) de cada uno de los centros seleccionados tanto de muy alta eficacia como de muy baja eficacia. Los centros que aparecen con su identificador subrayado muestran un interés añadido en el estudio, ya que son centros que obtienen un alto crecimiento en el promedio de sus residuos, pero, a su vez, muestran unos residuos extremos muy bajos y negativos.

6. CONCLUSIONES

Como conclusiones finales a este estudio, se hará especial mención a los objetivos, tanto generales como específicos, que se plantearon al inicio de la investigación. En primera instancia, para dar cumplimiento al objetivo principal de identificar y seleccionar los centros de muy alta y muy baja eficacia escolar de la Comunidad Autónoma Andaluza, se han empleado, como en un primer momento se consideró, los modelos jerárquicos lineales. Estos, han sido de gran utilidad para alcanzar, con éxito, la selección final de los centros y, además, para cumplir de igual manera el objetivo específico de conocer aquellas variables contextuales que influyen en el rendimiento final de los centros educativos. El análisis estadístico multinivel, bajo la perspectiva del estudio transversal contextualizado, ha sido un instrumento clave para controlar el efecto que las variables contextuales ejercen en los modelos y, por tanto, una herramienta necesaria para el cálculo del criterio final que se consideraría como identificador para la selección de centros de alta o baja eficacia: promedio de los residuos extremos.

De esta forma, se llegó al término de considerar centros muy eficaces aquellos cuyo promedio de los residuos en ambas competencias evaluadas (razonamiento matemático y comunicación lingüística) fueran muy altos y positivos. Además, se llegó a la conclusión de que las puntuaciones no logran ser un indicador válido de eficacia, ya que no se contempla la influencia de las covariables en el rendimiento o resultados finales. Incluso, en el caso de que las puntuaciones se establezcan mediante modelos jerárquicos lineales y se controle el efecto contextual, la correlación entre las puntuaciones y los residuos es alta (como se ha mostrado en el análisis de los datos) y prácticamente se presentan la misma información. Por ello, y a partir de la experiencia obtenida, lo más correcto ha sido proceder a la selección de los centros por su residuo y no en función de las puntuaciones directas.

Por otro lado, en relación a la decisión de incluir únicamente para el estudio transversal aquellas variables que tenían un carácter puramente contextual, sin duda, ha llegado a ser un indicador de veracidad, validez y objetividad a la hora de seleccionar los centros más y menos eficaces, ya que, para cada uno de ellos, se han considerado de igual forma las mismas cuestiones situacionales y su influencia en el aprendizaje final obtenido.

Las covariables que finalmente fueron incorporadas al estudio y destacadas en los resultados como factores que verdaderamente inciden en el rendimiento de los centros, muestran una gran variedad en función del año y competencia evaluada. Así, se puede llegar a pensar que la razón de alta eficacia en un centro educativo no se encuentra asociada a un patrón exacto y estandarizado de covariables que perdure en el tiempo. En la realidad educativa, se contemplan multitud de factores, y la verdadera razón de eficacia está en cada centro y depende de las covariables que puedan influir o cambiar en cada año. Esta es, una nueva visión propia de explorarse en un análisis cualitativo que ayude a conocer más en profundidad la realidad de cada centro en cuestión.

Por otro lado, comentar que, a diferencia de otros estudios de detección de centros eficaces, se planteó como objetivo general el de analizar los datos registrados de las evaluaciones diagnósticas contemplando una doble perspectiva basada en estudios de corte transversal contextualizado y longitudinal. Esta doble visión de análisis ha sido de gran ayuda y ofrece una calidad en los resultados obtenidos, ya que no se han considerado únicamente los centros en función de sus residuos extremos, sino que también se ha tenido en cuenta el crecimiento o (decrecimiento) de los residuos promedios para cada centro. De esta forma, la selección final de los centros más y menos eficaces, se ha realizado en función a dos criterios que han ayudado a dar una respuesta más real a la complejidad de la educación y su evaluación: promedio de los residuos extremos y promedio del crecimiento o decrecimiento de los residuos.

Como última objeción, decir que el presente trabajo es un válido estudio de detección de centros de muy alta y muy baja eficacia de la Comunidad Autónoma Andaluza, y que, puede ser utilizado para un análisis más en profundidad con la finalidad de obtener un guía de prácticas que ayuden y fomenten la mejora de la calidad de la educación.

7. LIMITACIONES

Como se sabe, los estudios enmarcados dentro del ámbito de la eficacia y la mejora escolar buscan conocer, como finalidad última, las prácticas o acciones que desempeñan los centros considerados como de muy alta eficacia para, poder diseñar planes que ayuden a renovar y enriquecer la calidad del sistema educativo en general. Para ello, es necesario utilizar una metodología de carácter cuantitativa basada en el análisis multinivel para la identificación de los centros y aquellas covariables que inciden en los resultados. A su vez, se recomienda combinarla con una metodología de tipo cualitativa que ayude a conocer la realidad que se da en esos centros. Teniendo en cuenta esta perspectiva, puedo hacer mención a una primera y fundamental limitación: el tiempo.

Para la elaboración de un estudio más válido, completo y, obtener unos resultados que ayuden verdaderamente a conseguir un avance y desarrollo en los planes de acción y enriquecimiento de la educación, hubiese sido de interés conocer el funcionamiento interno de los centros desde una perspectiva diferente. Es decir, adentrarse en la realidad de cada uno de estos y, comprender qué metodologías o actuaciones hacen que los centros educativos alcancen un buen rendimiento y logren en sus estudiantes, unos óptimos resultados en los aspectos evaluados. Desgraciadamente, llevar a cabo un estudio a este nivel era prácticamente imposible puesto que, investigaciones que contemplen una perspectiva tan amplia llevan años de realización. Solo en pensar en el desplazamiento y el tiempo que se necesita para llevar a cabo un estudio más en profundidad acerca de todos los centros seleccionados tanto de alta como de baja eficacia, hacía que básicamente ni se llegara a plantear la opción de un estudio que contemple ambas perspectivas. Aun así, no se descarta el seguimiento de este gran proyecto, en un futuro, para continuar investigando sobre este ámbito, ya que, hasta el momento, y con la realización de este trabajo, se ha obtenido un gran cimiento para comenzar a construir lo que podría ser un verdadero proyecto de mejora educativa en Andalucía.

Por otra parte, existen otras dos limitaciones claves en el estudio a las que se puede referenciar pero que, para mi sosiego, no se ha tenido un papel decisivo o de actuación ante ellas. Las limitaciones a las que hago mención, tienen origen en la información o los instrumentos con los que se ha contado para la realización de la investigación, es decir, las evaluaciones diagnósticas y los cuestionarios de familia proporcionados por la

AGAEVE. Siendo así, y como primer factor limitante en relación con el instrumento de análisis, lo encontramos en las covariables o variables contextuales del estudio.

En la mayoría de estudios de este nivel se observan como variables claves que influyen en el rendimiento final del alumnado y, consecuentemente, del centro educativo, los aspectos vinculados a la condición de ser inmigrante, la modalidad o método de enseñanza y el aprendizaje previo del alumnado del centro. Estas covariables, tanto a nivel de alumnos (nivel 1), como a nivel de escuelas (nivel 2) acaban teniendo un papel fundamental en los resultados obtenidos en los modelos multinivel. En general, la mayor parte de estos estudios llegan a coincidir en que las covariables señaladas ayudan a explicar en mayor proporción el modelo, ya que su influencia es relevante en el rendimiento final que alcanza el centro en cuestión.

Es lógico pensar que, la condición de ser inmigrante o la propia tasa de inmigración que presente un centro educativo, tenga correlación con la obtención de mejores o peores puntuaciones puesto que, solo por el factor idioma, se dificulta el aprendizaje en este sector del alumnado. En Andalucía, territorio implicado en el estudio, sabemos que es una razón muy actual y la tasa de inmigración es cada vez más abundante en provincias como Cádiz o Almería. Es una pena, que la Agencia Andaluza de Evaluación Educativa en ninguna de las evaluaciones realizadas hasta el momento, recoja en los cuestionarios de familia estos datos tan relevantes, puesto que en el caso de que se tuviera conciencia de la tasa de inmigración de los centros o las propias condiciones del alumno en cuanto a su conocimiento del idioma, se podía explicar el modelo más ampliamente y evitar el posible sesgo en el estudio. De forma análoga, los conocimientos previos del alumnado, la tasa de repetidores o la modalidad de enseñanza que presente el centro educativo, puede ser un importante factor explicativo en este tipo de estudios, a los cuales, no tenemos razón de influencia hasta ahora.

Por otro lado, en relación a un segundo aspecto a tratar como limitación para la investigación realizada y, que también hace referencia a los instrumentos utilizados para el estudio, es el hecho de la privacidad y seguridad de estos. Como he comentado en puntos anteriores, es lógico que, en un primer momento, la AGAEVE no permita una difusión abierta de los archivos de datos de las evaluaciones y establezca que su uso, tiene que estar restringido. Sin embargo, esto ha sido, en cierto punto, un obstáculo en la

investigación, ya que en todo momento se han trabajado con datos anónimos. Bajo esta perspectiva, hubiese sido interesante poder hacer un análisis más en profundidad acerca de cuáles son las provincias de la Comunidad Autónoma Andaluza que presenta centros con altos rendimientos. Estos hubiesen sido unos datos descriptivos solo a nivel territorial, y no bajo la finalidad de establecer un ranking de los mejores o peores centros. No obstante, no ha sido posible su realización debido a las limitaciones que la agencia ha ocasionado, y, por tanto, los resultados del estudio solo han podido dar a conocer el número de centros seleccionados, el por qué y las variables de carácter contextual que, a priori, y según los modelos jerárquicos realizados, afectan en el rendimiento y la eficacia escolar.

8. REFERENCIAS BIBLIOGRÁFICAS

Aguado, M. T. y Ballesteros, B. (2012). Presentación. Equidad y diversidad en la Educación Obligatoria. *Revista de Educación*, 358, 12-16.

Antillón, L. E., Larrazolo, N. y Backhoff, E. (2006). Igualación equipercenil del Examen de Habilidades y Conocimientos Básicos (EXHCOBA). *Revista Electrónica de Investigación y Evaluación Educativa (RELIEVE)*, 12(2), 205-217.

Blanco-Blanco, A., López-Martín, E., y Ruiz de Miguel, C. (2014). Aportaciones de los modelos jerárquico-lineales multivariados a la investigación educativa sobre el rendimiento. Un ejemplo con datos del alumnado español en PISA 2009. *Revista de Educación*, 365, 122-149. doi: 10.4438/1988-592X-RE-2014-365-267

Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models: applications and data analysis methods*. Newbury Park: Sage Publications.

Choi de Mendizábal, A., y Calero, J. (2013). Determinantes del riesgo de fracaso escolar en España en PISA-2009 y propuestas de reforma. *Revista De Educación*, 362, 562-593. doi: 10.4438/1988-592X-RE-2013-362-242

Creemers, B.P., Stoll, L. y Reezigt, G. (2007). Effective school improvement-ingredients for success: the results for an international comparative study of best practice case studies. En T. Townsend (Ed.), *International Handbook of school effectiveness and improvement* (pp. 825-838). Dordrecht, the Netherlands: Springer.

Cuttance, P. (1992). Evaluating the effectiveness schools. En D. Reynolds y P. Cuttance (Eds.), *School Effectiveness improvement. Research, policy and practice* (pp. 153- 169). Londres: Cassell.

De la Orden. A. y Jornet, J. M. (2012). La utilidad de las evaluaciones de sistemas educativos: el valor de la consideración del contexto. *Bordón*, 64(2), 69-88.

Field, A. (2013). *Discovering statistics using IBM SPSS statistics: and sex and drugs and rock'n' roll* (4.a edición). Los Angeles: Sage.

Fitz-Gibbon, C. T. (1997). *The Value-added National Project: Final Report. Feasibility studies for a national system of value-added indicators*. London: SCAA.

Gamazo, A., Martínez-Abad, F., Olmos-Migueláñez, S., y Rodríguez-Conde, M. J. (2018). Evaluación de factores relacionados con la eficacia escolar en PISA 2015. Un análisis multinivel. *Revista de Educación*, 379, 56-84. doi: 10.4438/1988-592XRE-2017-379-369.

Gamazo, A., Olmos-Migueláñez, S., y Martínez-Abad, F. (2016, Noviembre). Multilevel models for the assessment of school effectiveness using PISA scores. In *Proceedings of the Fourth International Conference on Technological Ecosystems for Enhancing Multiculturality* (pp. 1161-1166). ACM. doi:10.1145/3012430.3012663

González, A. M. (2015). *Dirección escolar exitosa en España. Un estudio de casos*. Universidad Autónoma de Madrid: España.

Hernández-Castilla, R., Murillo, F. J. y Martínez-Garrido, C. (2013). Factores de ineficacia escolar. REICE. *Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación (REICE)*, 12(1), 103-118.

Iñiguez-Berrozpe, T., y Marcaletti, F. (2018). Modelos lineales multinivel en SPSS y su aplicación en investigación educativa. *REIRE Revista d'Innovación y Recerca en Educació*, 11(1), 26-40. doi: 10.1344/reire2018.11.118984

Joaristi, L.; Lizasoain, L. y Azpillaga, V. (2014). Detección y caracterización de los centros escolares de alta eficacia de la Comunidad Autónoma del País Vasco mediante Modelos Transversales Contextualizados y Modelos Jerárquicos Lineales. *Estudios Sobre Educación*, 27, 37-61. doi: 10.15581/004.27

Lizasoain, L. y Joaristi, L. (2010). Estudio diferencial del rendimiento académico en lengua española de estudiantes de Educación Secundaria de Baja California (México). *Revista Iberoamericana de Evaluación Educativa*, 3(3), 116-134.

Lizasoain, L., Bereziartua, J. y Bartau, I. (2016). La formación permanente del profesorado en centros educativos de alta eficacia. *Bordón*, 68(2), 199-218. doi: 10.13042/Bordon.2016.68213

Mac-Gilchrist, B., Myers, K. y Reed, J. (2004). *The Intelligent School*. Londres: Sage.

Martínez-Abad, F., Chaparro, A. A., Lizasoain, L., Caso, J. y Urías, E. (2014). Identificación de escuelas de alto y bajo residuo, con base en puntajes de enlace y variables de contexto. UEE RT 14-003. México: Unidad de Evaluación Educativa.

Martínez-Abad, F. (2017). *Detección de buenas prácticas educativas en escuelas de alto valor añadido mediante técnicas de Big Data*. Instituto Universitario de Ciencias de la Educación, Universidad de Salamanca, España.

Martínez-Abad, F., Lizasoain, L., Castro, M., y Joaristi, L. M. (2017). Selección de escuelas de alta y baja eficacia en Baja California (México). *Revista Electrónica de Investigación Educativa*, 19(2), 38-53. doi: 10.24320/redie.2017.19.2.960

Martínez-Arias, R. (2009). Usos, aplicaciones y problemas de los modelos de valor añadido en educación. *Revista de Educación*, (348), 217-250.

Mata, P. y Ballesteros, B. (2012). Diversidad cultural, eficacia escolar y mejora de la escuela: encuentros y desencuentros. *Revista de Educación*, 358, 17-37. doi: 10.4438/1988-592X-RE-2012-358-181

Montero, R. (2016). *Modelos de regresión lineal múltiple*. Documentos de Trabajo en Economía Aplicada. Universidad de Granada. España.

Murillo, F.J. (2000). La investigación española sobre eficacia escolar. En A. Villa (Coord.), *Dirección Escolar y Organizaciones que aprenden* (pp. 1.002-1.020). Bilbao: Universidad de Deusto.

Murillo, F.J. (2005). *La investigación sobre Eficacia escolar*. Barcelona: Octaedro.

Perry, L., & McConney, A. (2010). School socio-economic composition and student outcomes in Australia: Implications for educational policy. *Australian Journal of Education*, 54(1), 72-85.

Reynolds, D., Hopkins, D., y Stoll, L. (1993). Linking school effectiveness knowledge and school improvement practice: Towards a synergy. *School Effectiveness and School Improvement*, 4(1), 37-58.

Santiago, K., Lukas, J. F., Joaristi, L., Lizasoain, L. y Moyano, N. (2008). A Longitudinal Study of Academic Achievement in Spanish: The Effect of Linguistic Models. *Language, Culture and Curriculum*, 21(1), 48-58.

Townsend, T., y Avalos, B. (2007). *International handbook of school effectiveness and improvement*. New York: Springer.