

El diseño experimental y la validación del currículo



Angel Vázquez Alonso
Inspección Técnica de Educación(*)

RESUMEN

En una coyuntura de gran vigencia del paradigma constructivista, coincidente en el tiempo con el proceso de reforma del sistema educativo, este artículo ofrece una reflexión sobre los aspectos metodológicos cruciales de la validación empírica de diseños curriculares, reflexión que puede servir a los autores de diseños curriculares para afinar sus propios métodos y técnicas de validación y suministrar a los profesores criterios adecuados para valorar críticamente y con mayor precisión la calidad y validez de diferentes currículos.

La última década ha sido fructífera en el avance de los conocimientos sobre la forma como aprenden los alumnos, y ello a pesar de no contar, todavía, con una gran teoría unificada del aprendizaje. El progreso en áreas concretas (detección de concepciones alternativas, resolución de problemas, clarificación de mecanismos, cambio conceptual y metodológico, estatus epistemológico ...) ha sido tan enorme y las líneas maestras parecen tan estabilizadas (constructivismo, psicología cognitiva -López Román, 1985- y modelos de procesamiento de la información) que el futuro no puede ser más prometedor. La contribución de los estudios realizados en torno al aprendizaje de las ciencias ha sido importante para el progreso en los conocimientos sobre el aprendizaje, en general. Desde la línea pionera iniciada por Piaget, y continuada en todo el mundo por numerosos seguidores,

al actual desarrollo de la investigación de las concepciones alternativas (Driver y Erickson, 1983; Driver, 1986), convertidas en piedra angular del paradigma constructivista (Bodner, 1986; Driver, 1988), se plantea el aprendizaje como un proceso activo y personal de construcción de significados y cambio conceptual (Hewson, 1981; Posner et al., 1983; Osborne y Wittrock, 1985; Coll, 1988) y metodológico (Gil, 1984; Carrascosa y Gil, 1985). Aunque no se ha llegado todavía a una teoría unificada del aprendizaje, todo el movimiento de investigación en torno al aprendizaje de las ciencias ha servido para consolidar definitivamente la Didáctica de las Ciencias como un área de investigación específica y con personalidad definida.

En el momento actual, fijados los referentes teóricos básicos, decantados de las investigaciones realizadas en años anteriores, y

(*) Servicio de Inspección Técnica de Educación
C/ San Miguel, 28 - 1º
07002 Palma de Mallorca (BALEARES)
Tfno. 971 / 71.61.70 - 71.61.76

asumidos por la comunidad científica y por los profesores, se inicia un movimiento dirigido al diseño de currículos educativos que ejemplifiquen y permitan la aplicación en el aula de estrategias de aprendizaje acordes con los principios epistemológicos y cognitivos del modelo constructivista del aprendizaje. Coincide, en nuestro país, con el momento de la implantación de la Reforma Educativa, cuyos ejes son el currículo abierto y la inspiración constructivista de la metodología del aprendizaje; el primero convierte a los profesores en diseñadores activos del proceso, y el segundo les requiere como investigadores en su propia aula. Esta coincidencia potenciará, sin duda, la tarea de creación de diseños curriculares en los momentos venideros, que ya es perceptible en el momento actual, hojeando algunas de las principales revistas. Lógicamente, este movimiento de diseño curricular e investigación en el aula, se aleja un poco de los lugares tradicionales de investigación pura, y es participado cada vez más por profesores en activo, movimientos de renovación y profesionales que ponen en práctica las propuestas de la investigación-acción (Stenhouse, 1984; Elliot, 1990) en las escuelas. Cada vez más, las revistas especializadas se hacen eco de propuestas concretas de aplicación curricular en la escuela, en ámbitos muy restringidos y para aprendizajes específicos, así como los grandes proyectos de investigación educativa institucional suelen reunir equipos multidisciplinares de especialistas, con suficientes apoyos y medios. Las iniciativas asumidas por los profesores suelen carecer de medios amplios, si se exceptúa el evidente entusiasmo de sus impulsores. En todos ellos es evidente la extraordinaria motivación y dedicación de sus profesores, pero la lectura de los informes de investigación publicados en revistas especializadas, pone de manifiesto, insistentemente, ciertas lagunas (principalmente metodológicas), que sugieren algunos comentarios para la reflexión, en orden a mejorar la calidad y amplitud de la validez de los proyectos curriculares diseñados y, por tanto, una evaluación más ajustada de cada uno de ellos.

El problema principal de la validación de proyectos curriculares nace de su propia génesis, ya que en la mayoría de los casos constituyen una creación personal de sus autores, circunstancia que provoca, a veces, el escepticismo en otros profesores, cuando el grado de subjetivismo es tal que impide una comprensión fácil; pero al mismo tiempo, nacen con la voluntad de ser útiles a otros, lo cual exige superar esta dificultad. El problema de validar correctamente un proyecto curricular está muy relacionado con la garantía de su difusión y generalización, para otros profesores y en otros contextos educativos, diferentes a aquéllos donde fue creado. Obviamente la trascendencia y generalización de un proyecto, que debe permitir una buena validación, no consiste en trasladar una mera copia de una escuela a otra, como si de la repetición del "librillo" se tratase, sino una auténtica y personal adaptación, mediatizada por la reflexión didáctica de cada profesor, en función de su contexto particular, dentro de la filosofía abierta y flexible del diseño curricular. Difícilmente un currículo con una validación deficiente o nula podrá trascender, en su aplicación práctica, más allá del círculo de su creación.

El objetivo de este artículo es comentar algunos aspectos metodológicos de la validación empírica de los diseños curriculares, como elemento de reflexión, de cara a promover diseños experimentales más válidos, que permitan definir con mayor precisión la calidad de un diseño, y, al mismo tiempo, potencien la aplicabilidad práctica de cada uno de ellos. También es necesario precisar que las observaciones que se harán se refieren siempre a diseños muy concretos, es decir, con una delimitación muy precisa de sus contenidos, que son los que habitualmente aparecen en las revistas especializadas, y no a proyectos curriculares institucionales, que, aunque sometidos a los mismos principios que se comentarán, tienen mejor garantizada su corrección por la abundancia de medios personales y materiales que congregan.

Validez

Es obvio que un currículo es un instrumento para ser aplicado en la práctica del aprendizaje de los alumnos, con unos propósitos educativos concretos que forman parte del propio diseño, pero que en general (en el modelo constructivista) se podrían resumir en la finalidad de producir aprendizajes realmente significativos. Es obvio también que el diseño curricular, en el plano meramente articulador de los distintos elementos que conforman el currículo (intenciones, contenidos, orientaciones didácticas, evaluación...), constituye, por sí misma, una actividad importante y meritoria, por el mero hecho de realizarse. Pero no puede perderse de vista el objetivo final: el currículo es un instrumento para conseguir un aprendizaje significativo. Inmediatamente la pregunta que se debe plantear (y responder) es: ¿este currículo concreto sirve para conseguir un aprendizaje significativo?, ¿este currículo concreto sirve para conseguir los objetivos que se propone? ¿sí? ¿no? ¿en qué grado? Esta es la pregunta fundamental de la validez: cualquier instrumento se dice que tiene validez (es válido) cuando sirve para aquello para lo que dice servir. Contratar la validez de un instrumento requiere poner a prueba la capacidad para alcanzar sus objetivos, y no otros.

Intuitivamente, sin tener conocimientos específicos sobre el concepto de validez, seguramente, todo el mundo se siente impelido a justificar las virtudes, la fundamentación y la coherencia de un proyecto curricular. Esto es una forma de iniciar (y aceptar) la necesidad e importancia del proceso de validación del currículo. Debería tenderse a conseguir que todos los artículos que recojan la aplicación de currículos de aprendizajes fundamenten lo más exhaustivamente posible su validez, o, al menos, expliciten claramente los elementos existentes en esta dirección. Y ello, no sólo porque es una exigencia metodológica en toda actividad de investigación, sino porque el importante alud de actividad, que presumiblemente se avecina en esta dirección, debe

permitir la discriminación clara de los elementos de validez aportados por unos u otros (o no aportados), y en definitiva, porque una buena validación es la mejor carta de presentación de cualquier proyecto curricular.

El desarrollo del concepto de validez ha tenido lugar principalmente en el marco de la medición psicológica y educativa. La American Psychological Association (APA) reconoce tres tipos de validez: validez de contenido, validez predictiva y de constructo. La validez de contenido va dirigida a justificar y fundamentar (generalmente en una teoría) los rasgos y características del currículo, es decir, se trata de demostrar la coherencia entre los elementos que conforman el currículo y las proposiciones de la teoría que lo sustenta. Por ejemplo, si se trata de diseñar un bloque curricular que remueva la tendencia de los alumnos a identificar la fuerza con el movimiento, debería justificarse que los elementos incluidos en el diseño están basados en las conclusiones de la investigación sobre este tópico, dentro del marco teórico adoptado. Los otros dos tipos de validez (predictiva y de constructo) están muy relacionados con los procesos de verificación empírica del currículo y las experiencias de su aplicación. La validación empírica debe ser un objetivo común y compartido por la mayoría de las propuestas de aplicación de diseños curriculares de aprendizaje pero, sin embargo, en muchos casos ni se plantea. Dentro del más puro estilo de la metodología científica, que los profesores de ciencias reivindicamos, es necesario ir hacia una exigencia de contraste empírico (contraste de hipótesis) de absolutamente todas las propuestas de diseño curricular. Debería comprenderse, en base a los propios postulados del constructivismo, que si un diseño curricular carece de una apoyatura empírica de resultados, aunque sea mínima, está gravemente cojo y se mantiene en el nivel de las propuestas teóricas, careciendo de soportes para justificar su validez práctica. En este aspecto, los grandes proyectos curriculares suelen cuidar estos aspectos metodológicos, gracias a la colaboración interdisciplinar de diver-

Los investigadores, y pueden servir de pauta metodológica para otros proyectos.

Dificultades metodológicas

Sin entrar en polémicas metateóricas sobre los planteamientos experimentales y suponiendo que la validación empírica pasa, de una u otra manera, por enfoques cualitativos y/o cuantitativos de corte empírico-experimental, las principales dificultades y errores metodológicos de la investigación educativa están relacionados, principalmente, con los siguientes aspectos: descripción y diseño general de la investigación, muestreo, procesos de medida e interpretación estadística.

El diseño de la investigación debe ir dirigido fundamentalmente a definir las variables relevantes y al control y manipulación adecuada de las mismas. Los problemas relacionados con la definición, control y manipulación de las variables suelen estar exacerbados, porque habitualmente no es posible una estricta aleatorización de las muestras. En consecuencia, debe cuidarse la precisión en la delimitación de cada una de estas operaciones.

Por otro lado, la mayoría de los diseños suelen emplear un diseño comparativo entre grupos de control y experimental: la descripción de los tratamientos que reciben ambos grupos no es suficientemente explícita de las variables o condiciones prácticas que marcan las diferencias entre los tratamientos. Es usual describir con énfasis el tratamiento experimental (el innovador) y casi nunca se describe el tratamiento que recibe el grupo de control. Asimismo, las descripciones se suelen detener en cuestiones formales y de segundo orden, pero se excluyen aspectos que pueden resultar decisivos para la validación, como por ejemplo, las características de los profesores de los grupos, tiempo de aplicación de la experiencia, materiales empleados por los alumnos, grado de insistencia en los conceptos que definirán las variables de contraste, etc. Cuando se comparan dos métodos de enseñanza/aprendizaje se pretende demostrar

que las diferencias observadas son debidas al método empleado. Si las numerosas variables que inciden en todo el proceso educativo no están cuidadosamente controladas y/o balanceadas entre los grupos, no se podrá garantizar que las diferencias observadas sean debidas, exclusivamente, al método de enseñanza que se somete a prueba, y no a otras variables o circunstancias espúreas. En consecuencia, pierde validez la comparación empírica y la generalización de los resultados.

El muestreo

Reconociendo la dificultad que tiene el muestreo, también es necesario afirmar que este tema recibe escasísima atención en la validación de proyectos curriculares. En general, los problemas más habituales del muestreo en educación están muy relacionados con las limitadas posibilidades de elección aleatoria de la muestra (generalmente los grupos están previamente constituídos), y en consecuencia, resulta gravemente afectada la representatividad muestral, y con ello, la generalización de resultados. Para mitigar estos efectos negativos se podrían intentar algunas estrategias, tales como describir la muestra en algunas otras de sus características, de modo que pueda justificarse la ausencia de sesgos importantes en la muestra, respecto al universo que representa, o también someter a contrabalanceo entre grupos las características que se estiman importantes para la validación buscada, etc.

En los diseños de grupo de control/experimental no sólo es importante la selección de los alumnos, sino también la de los profesores, que son los encargados de dirigir el proceso de enseñanza y aplicar la metodología; en particular, los profesores que atienden los grupos de control, y, concretamente, aclarar su intervención educativa en relación con los profesores de los grupos experimentales. ¿Tienen la misma capacidad y experiencia profesional? ¿Tratan el tema central de la investigación con la misma dedicación, con la

misma motivación, disponiendo de medios similares ...? ¿Conocen la investigación de la que forman parte?

La medida

En algún momento del proceso de validación es necesario realizar observaciones y/o medidas de las variables relevantes para el currículo que se somete a prueba. Sin duda, los problemas de medida son los más complejos del diseño experimental, por la multiplicidad de sus relaciones con diversos aspectos del diseño. Si tenemos en cuenta que los resultados de las medidas son la materia prima de los tratamientos estadísticos, es evidente que deben extremarse las precauciones; tres cuestiones son fundamentales: selección de las variables a medir, la ubicación temporal de las medidas y el instrumento de medida.

La selección de las variables que van a ser medidas supone elegir aquellas que se consideran más relevantes para el currículo que se trata de valorar. Normalmente, la teoría que sustenta el currículo ofrece pautas suficientes para su elección, dejando en manos del investigador/profesor la decisión última sobre su definición operacional en el contexto empírico donde se produce el estudio experimental. En este punto, el consejo más socorrido suele ser seleccionar variables que conformen un conjunto lo más amplio posible, o realizar diversos estudios con variables de distinto nivel de generalidad. El objetivo siempre es el mismo: conseguir la mayor representatividad de los resultados, mediante una descripción lo más amplia posible del estudio.

La ubicación temporal de las medidas es un tema que suele dejarse de lado, aceptando como general, implícitamente, el procedimiento test-retest, una medida antes del tratamiento de enseñanza y otra después.

Este procedimiento asume que las variables medidas mantienen una cierta constancia entre test y retest; sin embargo, una de las proposiciones básicas del constructivismo es la consecución de aprendizajes significativos,

que, entre otras cosas, suponen aprendizajes duraderos en el tiempo. Un diseño temporal, con diversas medidas repetidas en varios momentos podría validar mejor un currículo constructivista, enfatizando la importancia de la dimensión temporal/longitudinal en la evaluación de los aprendizajes, y por ende, la estabilidad de los mismos. Sin embargo, el instrumento empleado es el punto clave de una buena medida, con dos aspectos principales: su contenido y forma. El contenido de un instrumento debe responder a la operacionalización concreta de las variables seleccionadas como relevantes para el análisis, y una buena adecuación entre ambos es crucial para la validez de contenidos de todo el proceso. Una forma de asegurar esta adecuación se consigue, por ejemplo, a través de un proceso de ensayo previo y una buena experiencia en la elaboración del tipo de instrumento elegido.

La forma del instrumento se refiere al formato adoptado para sus items. Los tipos más usuales empleados con diseños grupales en educación son los tests de rendimiento, las escalas o cuestionarios y las parrillas de observación, siendo los dos primeros las formas más habituales empleadas en el tipo de validación curricular que nos ocupa. Desde el punto de vista metodológico, la principal diferencia entre ambos es la diferente precisión de las medidas que suministran: un test de rendimiento suele ser un instrumento válido y fiable, pues sus resultados están estandarizados y son independientes del observador (corrector), en tanto que los cuestionarios requieren la interpretación del observador, lo cual subjetiviza, en buena medida, los resultados. Por tanto, los tests de rendimiento o pruebas objetivas parecen la forma más precisa de realizar una medida fiable en educación, y de hecho, muchos estudios emplean este formato para realizar las medidas. Pero no es menos cierto que el mero empleo de un test no garantiza la fiabilidad en las medidas, ya que la mayoría de los estudios se quedan en la forma externa, sin conceder ninguna atención a la validez y fiabilidad del instru-

mento: en los informes de investigación no aparece discusión alguna sobre la validez y fiabilidad del instrumento, ni se contempla una etapa previa de validación y fiabilización, con una prueba piloto que permita una corrección de la misma antes de su aplicación definitiva, etc. El rasgo más significativo que define esta despreocupación general por la fiabilización del instrumento de medida es la sistemática aplicación de pruebas objetivas formadas por muy pocos ítems: una de las proposiciones elementales de la teoría clásica de los tests indica que la fiabilidad está muy relacionada con la longitud de los mismos, de manera que unos pocos ítems difícilmente llegarán a constituir un test fiable. La necesidad de producir pruebas de medida fiables (más largas) nos devuelve a la cuestión de la minuciosidad en la selección y operacionalización de las variables relevantes. Por último, debe indicarse que la fiabilidad es una condición necesaria para la validez, pero no suficiente; un test fiable puede ser válido, aunque no necesariamente, pero un test no fiable es imposible que sea válido.

La interpretación estadística

La fase final del diseño experimental requiere la traducción de los resultados al grado de validación alcanzado en el problema planteado, con dos operaciones fundamentales: la elección/aplicación de los criterios estadísticos y la valoración de los resultados obtenidos. Obviamente, si en los pasos previos se han cometido algunos de los errores metodológicos más habituales, al llegar a la fase final del diseño estadístico, las posibilidades de enderezar el proceso de validación son nulas. La selección del criterio estadístico, y su interpretación, debe ser adecuada y coherente con la muestra empleada y las medidas realizadas, y es una condición necesaria para la bondad de todo el proceso, pero no suficiente; la estadística, por sí misma, no permite hacer bueno un proceso de validación donde se han cometido errores en los pasos previos.

Específicamente, para esta última fase, el error más habitual que se aprecia es su ausencia: no se suele aplicar ninguna prueba estadística para verificar la relación entre variables, o la aceptación/rechazo de la hipótesis de nulidad para las diferencias comparadas entre grupos experimentales y de control, que es el planteamiento más habitual. En la mayoría de los casos, se limitan a exponer las tasas de aciertos y/o errores de cada uno de los grupos en las pruebas de medida o evaluación, dando por bueno que el grupo de control supere, con un cierto margen, los resultados del grupo experimental.

En los pocos casos que se utilizan estadísticos, está el problema general de emplear aquéllos que son inapropiados a la situación experimental, empobreciendo el poder de generalización de los resultados, y por ende, la validación del currículo que se pretende. Como es conocido, cada estadístico tiene unas condiciones de aplicación (paramétricos, no paramétricos, normalidad de la muestra, homoscedasticidad, etc.) que limitan su uso, invalidándolo en los casos que no se cumplen aquéllas; en consecuencia, la aplicación de un determinado estadístico requiere la comprobación previa de que se cumplen sus condiciones de aplicación en el caso concreto. La errónea selección estadística invalida absolutamente los resultados.

Por otro lado, cuando existe la posibilidad de seleccionar entre varios estadísticos debe tenerse en cuenta que no todos tienen la misma potencia, y tratar de elegir aquél que otorgue a los resultados una mayor significación (suele ser el más duro o contrario para la hipótesis). Generalmente, cuanto más potente es un estadístico, sus condiciones de aplicabilidad son más exigentes. En este proceso es también importante no caer en interpretaciones inadecuadas de los estadísticos usados; algunas de las más habituales suelen ser la consideración exclusiva de las medias (olvidándose de las desviaciones) en la descripción de las variables, deducir una relación causal entre variables a partir de una correlación significativa entre ellas, interpretar el resultado signi-

ficativo de un ANOVA como válido para todos los grupos (cuando hay más de dos), etc.

Por último, está el problema de la interpretación, es decir, traducir el alcance de los resultados estadísticos y las limitaciones del conjunto del proceso experimental de validación llevado a cabo (representatividad de la muestra, limitaciones del estadístico, grado de significación alcanzado, etc.), en relación con la(s) hipótesis planteada(s). Como en otros casos, el defecto más habitual es la ausencia de interpretación del conjunto de resultados, dando por supuesta una aceptación acrítica, y por tanto, una validación dogmática, opuesta a la metodología científica. Debe tenerse presente que los resultados numéricos, aisladamente por sí mismos, no indican gran cosa en un problema tan complejo como la validación curricular; es necesario someterles a un proceso crítico de reflexión y discusión, a la luz del cuerpo de conocimientos científicos que ha inspirado la investigación, considerando especialmente los resultados obtenidos en proyectos e investigaciones similares.

Como colofón, es necesario reiterar que estos comentarios metodológicos sobre la validación experimental de currículos escolares no serían posibles sin la referencia de una actividad que comienza a ser notable en este terreno. Por tanto, este artículo representa un reconocimiento implícito al esfuerzo realizado por muchos profesores en esta dirección, y va guiado por el deseo de apoyar y mejorar los aspectos metodológicos de esta actividad investigadora en la escuela. No se incluye una bibliografía especializada sobre el tema, porque ésta es tan amplia, afortunadamente, que cualquiera de los numerosos manuales existentes puede servir para una buena orientación. Tampoco se han incluido referencias a informes de validación ya publicados por razones obvias, y porque el objetivo de estas líneas es simplemente ayudar a reflexionar sobre los aspectos metodológicos, y no generar polémicas.

REFERENCIAS

- BODNER, G. M. (1986). Constructivism: A Theory of Knowledge. *Journal of Chemical Education*, 63 (10), pp. 873-878.
- CARRASCOSA ALIS, J. y GIL PÉREZ, D. (1985). La "metodología de la superficialitat" i l'aprenentatge de les ciències. *Ensenyanza de las ciencias*, 3, pp. 113-120.
- COLL, C. (1988). Significado y sentido en el aprendizaje escolar. Reflexiones en torno al concepto de aprendizaje significativo. *Infancia y Aprendizaje*, 41, pp. 131-142.
- DRIVER, R. (1986). Psicología cognoscitiva y esquemas conceptuales de los alumnos. *Ensenyanza de las Ciencias*, 4, pp. 3-15.
- DRIVER, R. y ERICKSON, G. (1983). Theories in action: Some theoretical and empirical issues in the study of students' conceptual frameworks in science. *Studies in Science Education*, 10, pp. 37-60.
- DRIVER, R. (1988). Un enfoque constructivista para el desarrollo del currículum en ciencias. *Ensenyanza de las ciencias*, 6, pp. 109-120.
- DRIVER, R.; GUESNE, E. y TIBERGHIE, A. (1989). *Ideas científicas en la infancia y en la adolescencia*. Madrid, Morata-M.E.C. (traducción de *Children's ideas in science*, Open University Press, 1985).
- ELLIOT, J. (1990). *Investigación-acción en Educación*. Morata, Madrid.
- GIL PÉREZ, D. (1984). La metodología científica y la enseñanza de las ciencias: unas relaciones controvertidas. *Ensenyanza de las Ciencias*, 4 (2), pp. 111-121.
- LÓPEZ ROMÁN, J. (1985). Psicología cognitiva y educación. *Bordón*, 259, pp. 557-570.
- HEWSON, P. W. (1981). A conceptual Change Approach to Learning Science. *European Journal of Science Education*, 3, pp. 383-396.
- OSBORNE, R. y WITTRICK, M. (1985). The genepathic learning model and its implications for science education. *Studies in Science Education*, 12, pp. 59-87.
- POSNER, J. G.; STRIKE, K. A.; HEWSON, P. W. y GERTZOG, W. A. (1982). Accommodation of a scientific conception: toward a theory of conceptual change. *Science Education*, 66, pp. 211-227.
- STENHOUSE, L. (1984). *Investigación y desarrollo del currículum*. Morata, Madrid.

SUMMARY

In a moment when the constructivist paradigm is highly operative, coinciding in time with the process of the educational system reform, this paper reflects on the crucial methodological aspects of the empiric validation of curricular designs, a reflection wich can be used by the authors of curricular designs to refine their own methods and validation techniques and to provide teachers with appropriate criteria to assess critically and more precisely the quality and validity of different curricula.

RÉSUMÉ

Maitenant, quand le paradigme constructiviste et de grande actualité, coïncidant dans le temps avec le procès de réforme du système éducatif, cet article réfléchit aux aspects méthodologiques fondamentaux de la validation empirique des dessins curriculaires, réflexion qui peut servir aux auteurs de dessins curriculaires pour améliorer leurs propres méthodes et techniques de validation et donner aux professeurs des critères appropriés pour évaluer de façon critique et avec plus de précision la qualité et la validité de différents curricula.