

En búsqueda de la explicación de los resultados educativos: posibilidades y limitaciones de los estudios internacionales de evaluación

Alejandro Tiana Ferrer

Resumen: *Los estudios internacionales de evaluación del rendimiento educativo, que se comenzaron a desarrollar a mediados del siglo XX y han experimentado un auge considerable en las dos últimas décadas, han perseguido un doble propósito. Por una parte, sustentar la comparación rigurosa de los resultados obtenidos por los estudiantes en distintos sistemas educativos, ofreciendo como resultado un valioso conjunto de datos que han contribuido a mejorar nuestro conocimiento acerca del rendimiento educativo. Por otra, aportar elementos para explicar los resultados obtenidos, en función de las variables que inciden en ellos.*

En este trabajo se analiza el segundo de dichos propósitos, estudiando las posibilidades que dichos estudios abren y sus aportaciones a tal explicación, así como las dificultades encontradas y las limitaciones que ofrecen para explicar el efecto de los diferentes factores identificados. Para ello se distinguen dos grandes tipos de variables, unas denominadas extrínsecas, relacionadas con factores externos al sistema educativo, y otras intrínsecas, relativas a su organización y funcionamiento. Dada la perspectiva macroscópica adoptada en tales estudios, se concluye que la información que aportan a dicha explicación resulta forzosamente limitada, debiendo complementarse por otro tipo de estudios de investigación y evaluación capaces de permitir análisis más detallados.

Palabras clave: *Educación comparada, evaluación del rendimiento, evaluación a gran escala, índice socioeconómico y cultural (ISEC), eficacia escolar.*

INTRODUCCIÓN

Desde la última década del siglo XX asistimos a una preocupación creciente por la mejora de la calidad de la educación, tanto en los países desarrollados como en los que están en vías de serlo. Factores tales como la presión económica internacional, la expansión del proceso de globalización, la introducción masiva de las nuevas tecno-

logías de la información y la comunicación o la necesidad acuciante de introducir cambios en los modelos productivos en la actual época de crisis económica han vuelto a atraer la atención hacia la educación, considerada nuevamente como un factor decisivo para asegurar el desarrollo de los pueblos.

La celebración en 1990 de una relevante reunión de los ministros de Educación de los países miembros de la *Organisation for Economic Co-operation and Development* (OECD, en castellano OCDE), con el lema de *Una educación y una formación de calidad para todos* (OECD, 1992), supuso la puesta en marcha de numerosas iniciativas con ese propósito, que desde entonces no han cesado de extenderse. Aunque no se puede ignorar que el término *calidad* es polisémico e impreciso, se ha convertido en uno de los argumentos clave (incluso en un *fetiché*, podría decirse) de las políticas educativas contemporáneas (Tiana, 2006).

En este contexto de renovada preocupación por la educación y su calidad, se ha difundido durante las dos últimas décadas el interés por evaluar y comparar la situación y los resultados de los sistemas educativos y han aparecido nuevas iniciativas en ese sentido, entre las que destaca la construcción de sistemas de indicadores de la educación y la realización de ambiciosos estudios internacionales de evaluación del rendimiento educativo. Una de las manifestaciones más patentes del avance registrado se encuentra en el impacto producido por la publicación de los resultados de los sucesivos ciclos del proyecto *Program for International Student Assessment* (PISA), que ha suscitado importantes debates públicos en muchos y muy diversos países (Kotthoff y Pereyra, 2009).

1. COMPARAR Y EXPLICAR LOS RESULTADOS DE LA EDUCACIÓN MEDIANTE ESTUDIOS INTERNACIONALES

En realidad, el desarrollo de esos estudios representa la consolidación de una tendencia anterior, que no se puede considerar estrictamente novedosa. En efecto, hay que recordar que sus antecedentes se remontan al final de los años cincuenta, cuando un grupo de reconocidos investigadores en educación decidió crear una asociación internacional para evaluar los resultados de los sistemas educativos. Así es como se constituyó la International Association for the Evaluation of Educational Achievement (IEA), tras una primera reunión celebrada en 1958 en el Instituto de Educación de la UNESCO en Hamburgo (Tiana, 2001).

La propuesta de aquellos investigadores consistía en planificar y llevar a cabo estudios internacionales del rendimiento educativo adoptando una perspectiva comparada, lo que resultaba realmente novedoso para su tiempo. Desde su punto de vista, los estudios comparativos de los sistemas educativos debían tomar necesariamente en consideración los resultados que conseguían los estudiantes, cosa que no era habitual hasta entonces. Aunque no menospreciaban los análisis basados en variables tales como los recursos disponibles, los procesos de escolarización o las con-

diciones en que se desarrolla la enseñanza, consideraban que la comparación de los sistemas educativos debía completarse con la emisión de juicios rigurosos sobre los resultados alcanzados por los alumnos.

No obstante, aquel grupo de investigadores que constituyó la IEA no quería limitarse a medir y comparar los resultados educativos conseguidos por los estudiantes, sino que pretendía explicarlos de manera convincente. Además de calcular y valorar los resultados, pretendían identificar los principales factores capaces de explicar las diferencias de rendimiento encontradas y analizar de qué modo interactuaban entre sí. Esa explicación constituía, en última instancia, la base para diseñar cualquier acción de mejora de la educación, que era lo que realmente les preocupaba.

En ese contexto surgió la idea de considerar el mundo como un *laboratorio educativo* (Purves, 1993), concepción ciertamente sencilla, pero cuya originalidad hay que reconocer. Su fundamento se encuentra en la explotación de la *variación natural* que existe entre los sistemas educativos. Frente al caso de otros campos científicos, que permiten la aplicación de procedimientos experimentales para estudiar el efecto de determinadas intervenciones, en el ámbito educativo es muy complicado adoptar modelos semejantes e incluso para muchos resulta éticamente discutible. Los investigadores suelen recurrir a otro tipo de métodos, en muchos casos empíricos pero pocas veces de carácter estrictamente experimental. Para salvar esa limitación, los fundadores de la IEA consideraron que podría explorarse con fines comparativos el efecto de las diferencias *naturales* existentes entre los sistemas educativos (aunque, en realidad, puede argumentarse que dichas diferencias derivan en buena medida de su historia y su contexto).

De acuerdo con esa concepción, los sistemas educativos deben ser considerados como conglomerados complejos, compuestos por un número elevado de escuelas, clases, profesores, estudiantes y comunidades. Cada sistema comparte rasgos comunes con otros, pero también presenta algunas características propias. Por lo tanto, sería teóricamente posible comparar los resultados conseguidos por los estudiantes de distintos sistemas educativos, controlando tales variables diferenciales que inciden en el rendimiento. Ese tipo de comparación permitiría analizar y valorar la influencia de dichos factores, contribuyendo a explicar sus efectos respectivos. Así, por ejemplo, podrían valorarse las consecuencias de los diversos modos de diseño y desarrollo del currículo, la integración o segregación de los alumnos en vías académicas alternativas, los modelos de formación del profesorado o la interacción didáctica en el interior del aula, por no citar sino algunos factores frecuentemente estudiados. Sin necesidad de introducir cambios en un sistema educativo determinado, lo que no siempre resulta posible ni conveniente, la existencia de esa *variación natural* permitiría analizar y valorar el efecto de los principales factores asociados con el rendimiento que no se presentan de igual modo en todos los sistemas. Esa es, en síntesis, la concepción del *mundo como laboratorio educativo* que la IEA difundió entre los investigadores de la educación.

Esa idea directriz sirvió de esquema básico para diseñar y poner en marcha los primeros estudios emprendidos por la IEA, entre los que destacan el primer Estudio de Viabilidad (*Feasibility Study*), de carácter piloto y desarrollado entre 1959 y 1962 en doce sistemas educativos¹; el Primer Estudio Internacional de Matemáticas (*First International Mathematics Study – FIMS*), que se llevó a cabo entre 1959 y 1967; el Estudio de Seis Materias² (*Six-subject Survey*), iniciado en 1966 y finalizado a comienzos de los setenta; y el Segundo Estudio Internacional de Matemáticas (*Second International Mathematics Study – SIMS*), realizado entre 1964 y 1967. Posteriormente, durante los años setenta y ochenta, la IEA desarrolló el Estudio del Entorno Escolar (*Classroom Environment Study*, 1978-1981), el Estudio Internacional de Composición Escrita (*International Study of Achievement in Written Composition*, 1980-1986), y el Segundo Estudio Internacional en Ciencias (*Second International Science Study – SISS*, 1980-1987). Ya en los noventa desarrollaría el estudio sobre Educación Preescolar (*Pre-Primary Education*) e iniciaría los ciclos conocidos respectivamente como Tercer Estudio Internacional de Matemáticas y Ciencias (*Third International Mathematics and Science Study – TIMSS*), Comprensión Lectora (*Reading Literacy Study* y *Progress in Reading Literacy Study – PIRLS*), Tecnologías de la Información en Educación (*Computers in Education* y *Second Information Technologies in Education Study – SITES*), y Educación Cívica (*Civic Education*), que con distintas características y vicisitudes continúan desarrollándose en la actualidad.

Este conjunto de trabajos, que ha dado lugar a una amplia serie de informes y publicaciones (Degenhart, 1990; Papanastasiou, 2004), ha proporcionado una valiosa información sobre el nivel comparativo de rendimiento de los sistemas educativos, por una parte, y sobre las variables asociadas a los resultados, por otra. Para permitir tales logros, la IEA ha aplicado una metodología generalmente cuantitativa, basada, por una parte, en la elaboración de marcos conceptuales rigurosos sustentados en análisis curriculares de las etapas estudiadas y, por otra, en la aplicación de pruebas estandarizadas de rendimiento y cuestionarios dirigidos a profesores, directores escolares y alumnos, construidos cooperativamente con el propósito de conseguir que estén libres de sesgos culturales, idiomáticos o de otro tipo. En ocasiones se han aplicado otros procedimientos, tales como realización de tareas prácticas (*performance tasks*) o técnicas cualitativas (estudios de caso, por ejemplo), pero insistiendo siempre en su carácter estandarizado. La recogida de datos por medio de dichos instrumentos permite aplicar técnicas de análisis rigurosas, respetuosas de las condiciones de una comparación fiable y válida. Con el paso de los años, los procedimientos concretos de recogida y análisis de datos han ido variando, esforzándose siempre por mantener su carácter innovador.

1 La IEA es una asociación voluntaria e independiente, no gubernamental, compuesta por instituciones de investigación que representan a sistemas educativos diferentes. Esa definición hace que países como Bélgica o el Reino Unido tengan una doble representación (Bélgica francesa y flamenca o Inglaterra-Gales y Escocia), aunque la mayoría de sus miembros correspondan a países con un único sistema educativo (incluso en el caso de países federales como Alemania, Estados Unidos, Australia o Canadá). Para más información sobre la IEA y los estudios que desarrolla en la actualidad, se puede consultar su página web (www.iea.nl).

2 Comprendía las materias de Ciencias, Comprensión lectora, Literatura, Francés e Inglés como lenguas extranjeras y Educación Cívica.

Con posterioridad, otros organismos internacionales comenzaron a producir sus propios estudios de rendimiento educativo, aplicando este tipo de procedimientos y logrando avances importantes, tanto conceptuales como metodológicos. Entre ellos hay que destacar especialmente los estudios promovidos por la (OECD). Aunque su interés por los asuntos educativos viene de antiguo (Papadopoulos, 1994), en la década de los noventa puso en marcha un ambicioso proyecto de construcción de indicadores internacionales de la educación, que ha dado importantes frutos (CERI, 1994). Se trata del denominado Proyecto INES, que ha producido la serie de informes que viene publicándose desde 1992 con el título genérico de *Education at a Glance / Regards sur l'éducation*.

Los indicadores elaborados por la OECD se refieren tanto al contexto de la educación como a los recursos humanos y económicos utilizados, los procesos escolares y los resultados logrados. No obstante, esta última dimensión planteó dificultades, dado que los únicos datos existentes en los años noventa eran los procedentes de los estudios de la IEA, que no siempre se ajustaban a las demandas formuladas, ni por el número de países participantes, ni por las áreas evaluadas, ni por su periodicidad. Para salvar esa carencia de indicadores de resultados educativos, la OECD diseñó y puso en marcha a finales de esa década el proyecto denominado PISA, que pretende valorar el grado de formación de los jóvenes de quince años, en tres áreas básicas: matemáticas, lectura y ciencias. La primera recogida de datos se realizó en el año 2000 (OECD, 2001b), repitiéndose a partir de entonces cada tres años, aunque haciendo énfasis sucesivamente en cada una de las áreas mencionadas (OECD, 2004 y 2007).

La OECD también puso en marcha otros estudios orientados a la evaluación de las capacidades de la población adulta (*International Adult Literacy Survey, International Life Skills Survey*), que han producido algunos resultados de interés (OECD & Statistics Canada, 1995 y 2000). Su objetivo consiste en aproximarse a una estimación del capital humano de los países, a través del grado de formación de su población en edad laboral (de 16 a 65 años).

Por su parte, la UNESCO ha emprendido algún proyecto de la misma naturaleza, entre los que destaca especialmente el denominado Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación (LLECE), cuyos trabajos se han desarrollado también en la última década (OREALC, 1998 y 2008).

Ahora bien, toda esta serie de estudios, aun siendo semejantes, tienen sin embargo características distintas, que conviene destacar. Ante todo, hay que distinguir los proyectos directamente orientados a evaluar los resultados conseguidos por los estudiantes (como son el TIMSS, PIRLS o PISA) de aquellos otros que tienen unos objetivos más amplios, como es el caso de la construcción de sistemas de indicadores (Proyecto INES). Si bien es cierto que estos últimos suelen incluir algunos indicadores de resultados, no siempre constituyen su núcleo central, ni consiguen en todos los casos aportar datos originales. Puede decirse que los sistemas de indicadores de la educación elaborados hasta el momento son tributarios del desarrollo experimentado por la estadística educativa, más que de los avances registrados en el campo de la evaluación del rendimiento.

Descendiendo al análisis de los estudios propiamente dichos de evaluación del rendimiento educativo, conviene distinguir al menos cuatro aspectos diferenciales, cada uno de los cuales puede presentarse en forma dicotómica, como un eje lineal con dos extremos, aunque puedan identificarse posiciones intermedias.

- a) Una primera distinción es la que puede establecerse entre los estudios encaminados a la *evaluación del rendimiento escolar* y los orientados a la *estimación del nivel formativo* de una población. Los del primer tipo pretenden evaluar los resultados que logran los alumnos que alcanzan un determinado grado o curso escolar. Es el caso de estudios como TIMSS, centrado en los aprendizajes adquiridos en matemáticas y ciencias por los alumnos de los cursos cuarto y octavo (equivalentes a los diez y catorce años de edad modal) o PIRLS, centrado en el aprendizaje lector logrado al alcanzar el cuarto curso de escolaridad. En ambos estudios, se evalúa a los alumnos matriculados en dicho curso, con independencia de su edad, aunque esa variable se controle y se tenga en cuenta en los análisis. Los estudios del segundo tipo pretenden estimar el nivel formativo alcanzado por las personas que tienen una determinada edad, independientemente de su ubicación escolar. Es el caso de PISA, que evalúa las competencias adquiridas por los jóvenes de quince años de edad en relación con las matemáticas, la lectura y las ciencias, o de IALS o su sucesor PIAAC (Programa para la Evaluación Internacional de las Competencias de los Adultos), cuya población de estudio es la comprendida entre 16 y 65 años. En estos últimos casos, los sujetos evaluados son elegidos al azar entre quienes tienen dicha edad o edades, con independencia del nivel escolar que hayan alcanzado. Mientras que los estudios del primer tipo están más directamente relacionados con la eficacia interna del sistema educativo, los del segundo se orientan a valorar sus resultados finales.

- b) Una segunda distinción es la existente entre los estudios centrados en las *áreas curriculares básicas* y los dedicados a la evaluación de *capacidades o competencias transversales*. Los del primer tipo pretenden medir el grado de dominio de diversas habilidades ligadas a áreas tales como las matemáticas, la lengua materna o extranjera o las ciencias, esto es, a áreas instrumentales básicas del currículo escolar. Es el caso de TIMSS, PIRLS o PISA, antes mencionados. Los estudios del segundo tipo pretenden medir el grado de desarrollo de diversas capacidades no vinculadas a áreas curriculares concretas, sino relacionadas con otras dimensiones formativas, como pueden ser la educación cívica o la utilización de las nuevas tecnologías de la información. Es el caso de los estudios CIVICS o ICCS de la IEA, que evalúan los conocimientos y las actitudes cívicas de los jóvenes y adolescentes, y también del SITES o el componente adicional del PISA, que valoran la familiaridad con la tecnología informática y el uso que los escolares hacen de la misma. Mientras que los estudios del primer tipo están muy relacionados con el currículo escolar, los segundos tienen que ver con un tipo de competencias más amplias y vinculadas con distintos tipos de aprendizajes.

- c) Una tercera distinción es la que contrapone los estudios centrados en la evaluación de las *capacidades o conocimientos adquiridos por los estudiantes individualmente considerados* con los que tienen como objetivo estimar el *rendimiento educativo general de la población escolar*. Los del primer tipo se basan en la aplicación de pruebas idénticas para todos los alumnos evaluados, lo que proporciona la posibilidad de establecer comparaciones o de realizar análisis a partir de los resultados individuales. Su contrapartida consiste en que obligan a limitar la cobertura curricular de las pruebas, para evitar una extensión excesiva, lo que implica reducir la profundidad de los análisis subsiguientes. Los estudios del segundo tipo se basan en la aplicación de pruebas rotadas, respondiendo los alumnos a combinaciones diferentes de cuadernillos parciales, con lo que no existe un modelo único de prueba. A cambio, puede ampliarse notablemente la cobertura curricular y se pueden hacer análisis más profundos, aunque sin poder establecer conclusiones de nivel individual, sino únicamente grupal. Cada tipo de estudios permite llegar a conclusiones de distinta naturaleza sobre la situación del sistema educativo, de los centros docentes y de los estudiantes.
- d) Una cuarta distinción puede establecerse entre los estudios que evalúan los *resultados educativos en un momento concreto* y los que pretenden medir el *progreso educativo a lo largo del tiempo*. Los del primer tipo proporcionan una información valiosa acerca del estado de una parcela determinada del sistema educativo en un momento dado, en una circunstancia temporal concreta. Los del segundo tipo se proponen valorar el progreso producido en un lapso de tiempo determinado. Los estudios más sofisticados de los que se desarrollan en la actualidad combinan ambos propósitos. Así, por ejemplo, TIMSS, PIRLS y PISA están concebidos como ciclos continuados, lo que permite, por una parte, conocer cuál es la situación educativa en un año concreto, y por otra, valorar el progreso producido cada varios años. De ese modo, los países participantes no sólo pueden compararse con otros en un momento dado, sino que pueden analizar cuál es su evolución.

Como puede apreciarse, los estudios internacionales tienen características diferentes, que permiten establecer distinciones entre ellos, tanto por el enfoque que adoptan como por la población elegida y el tipo de comparación predominante. Por lo tanto, se puede concluir que para poder interpretar o valorar correctamente un estudio determinado hay que analizar y tener en cuenta sus características específicas.

Los estudios mencionados tienen dos objetivos fundamentales, que no deben entenderse necesariamente como contrapuestos. Por una parte, se proponen presentar una información confiable sobre la situación comparativa de los diversos sistemas educativos, o al menos sobre algunas parcelas concretas de los mismos. Por otra

parte, pretenden ofrecer algunas claves para explicar e interpretar los resultados logrados. Aunque ambos objetivos no son alternativos, pues muchos estudios combinan ambos, cabe distinguirlos a efectos de análisis.

Esa voluntad de explicación que subyace en muchos estudios está ligada a un compromiso con la mejora de la educación, pues para poder emprender acciones de mejora es necesario conocer cuáles son las claves fundamentales del rendimiento educativo. Sin embargo, hablando en términos rigurosos, existe la convicción generalizada de que no es sencillo elaborar modelos causales del rendimiento, pese a los intentos que se han realizado en ese sentido (Álvaro et al., 1990). Por ese motivo, y consciente de esa dificultad, me he limitado aquí a presentar y analizar los factores que se toman generalmente en consideración a la hora de explicar los resultados, sin pretender llegar a conclusiones definitivas.

Desde esa perspectiva, las variables que pueden ayudarnos a explicar los resultados de la educación son muy diversas. No obstante, en ese amplio conjunto se pueden distinguir dos grandes categorías:

- Un primer grupo está compuesto por variables que se refieren a factores que no están directamente ligados al funcionamiento del sistema educativo o de los centros escolares. Guardan relación con las condiciones sociales y económicas en que se desenvuelve la tarea educativa, por lo cual no resulta sencillo actuar sobre ellas desde el marco estrictamente escolar. Aquí se denominan *variables extrínsecas*.
- Un segundo grupo está formado por variables que se refieren a factores directamente relacionados con el funcionamiento del sistema educativo o de los centros escolares. Tienen que ver con el modo de organización del sistema educativo, así como con el funcionamiento de las escuelas y el desarrollo de los procesos de enseñanza y aprendizaje, por lo que puede actuar-se sobre ellas desde el marco educativo y escolar. Aquí se denominan *variables intrínsecas*.

Entre la diversidad de variables de ambos tipos que pueden identificarse, los investigadores y evaluadores se han centrado especialmente en algunas de ellas, que he sintetizado en el cuadro siguiente y que servirán de eje de análisis en los apartados sucesivos.

Variables extrínsecas	Variables intrínsecas
<ul style="list-style-type: none"> • Nivel socioeconómico y cultural de los estudiantes • Nivel de desarrollo educativo y cultural de la población • Recursos destinados a la educación 	<ul style="list-style-type: none"> • Organización del sistema educativo • Procesos de centro escolar • Procesos de aula

2. BUSCANDO EXPLICACIONES: VARIABLES EXTRÍNECAS DEL RENDIMIENTO

Entre las variables extrínsecas, hay tres que tienen especial interés, además de contar con tradición en la investigación educativa. Se trata del nivel socioeconómico y cultural de los estudiantes y de sus familias, el nivel de desarrollo educativo y cultural del conjunto de la población y los recursos que se destinan a la educación. Aunque esa relación no agota el conjunto de variables extrínsecas que podrían considerarse, tiene al menos en cuenta las más relevantes.

— Nivel socioeconómico y cultural de los estudiantes

A la hora de explicar los resultados logrados por los estudiantes, tomar en consideración el nivel socioeconómico y cultural de la población evaluada constituye una práctica habitual de la investigación educativa (Tiana, 2002). Hoy se acepta plenamente que los estudiantes están condicionados por una serie de factores de carácter restringido o *microsocial*, relacionados con las características socioeconómicas y culturales concretas de su grupo familiar y de su entorno cercano, y por otros de carácter más amplio o *macrosocial*, relacionados con aspectos tales como su posición social o las características de su entorno más distante. En consecuencia, hay que tomar en consideración ambos tipos de variables cuando se analizan los resultados logrados por los alumnos.

Puesto que en el párrafo anterior se habla del nivel socioeconómico y cultural, resulta necesario señalar que los primeros estudios internacionales de evaluación del rendimiento que se llevaron a cabo hace ya medio siglo optaron generalmente por controlar el nivel estrictamente socioeconómico de la población analizada, adoptando un enfoque por entonces habitual en los estudios empíricos de sociología de la educación. Sin embargo, posteriores análisis más detallados pusieron de manifiesto que los resultados educativos tienen mayor relación con los indicadores de posición social cuando éstos incluyen algunas variables relativas al grado de riqueza cultural del entorno.

Ese fenómeno ha sido bien analizado en algunos estudios de evaluación de ámbito nacional. Así, por ejemplo, en la evaluación de la entonces nueva Educación Primaria que se llevó a cabo en España en 1995³ se aprovechó la ocasión para construir dos índices complementarios y parcialmente coincidentes, a partir de algunas variables extrínsecas como las que aquí se mencionan. El primero fue el *índice socioeconómico*, compuesto por variables relativas a la profesión de los padres y la posesión familiar de bienes materiales. El segundo fue el *índice sociocultural*, compuesto

3 Ese estudio fue el primero íntegramente diseñado y desarrollado por el Instituto Nacional de Calidad y Evaluación (INCE). En consecuencia, se utilizó para hacer varios ensayos pilotos, como algunos que se comentan en estas páginas.

por variables relativas a los estudios y la profesión de los padres, más otras referidas a las expectativas educativas para los hijos y el acceso a algunos bienes culturales básicos (como libros, diccionarios u ordenador). Los análisis estadísticos realizados demostraron que la relación del segundo índice con los resultados logrados por los alumnos era claramente mayor que la del primero, lo que avala la importancia que poseen y el carácter predictivo que tienen las variables de tipo educativo y cultural por encima de las estrictamente económicas (INCE, 1996). Dicho de otro modo, el grado de riqueza material de la familia parece influir menos en los resultados académicos de los hijos que la riqueza del entorno educativo y cultural, lo que además concuerda con nuestra experiencia cotidiana.

En aplicación de esas conclusiones, los cuestionarios para las familias o los alumnos que acompañan a los recientes estudios españoles de evaluación (como los estudios periódicos de evaluación de la Educación Primaria o los relativos a la Educación Secundaria Obligatoria), al igual que los del proyecto PISA, incluyen preguntas acerca de aspectos tales como el número de libros disponibles en el hogar, la posesión de ordenador o diccionarios, el nivel educativo de los padres o las expectativas académicas hacia los hijos. La combinación de estas variables culturales con otras más clásicas como los bienes materiales en el hogar o la situación laboral de los progenitores eleva considerablemente el porcentaje explicado de la varianza, por lo que su uso se ha convertido en una práctica habitual.

En consecuencia, los estudios recientes de evaluación han incluido la elaboración de índices socioeconómicos y culturales con el propósito de contribuir a explicar las diferencias de resultados educativos. Uno de los modelos más completos actualmente en vigor es el Índice Socioeconómico y Cultural (ISEC) calculado por PISA, que permite llevar a cabo interesantes análisis acerca de la equidad de la educación, ya que permite estimar la correlación que existe entre el origen y la posición social de los estudiantes y sus resultados escolares (OECD, 2001b, 2004 y 2007).

Esos índices son utilizados para ponderar los resultados obtenidos, eliminando el efecto de las diferencias sociales y estimando con mayor precisión el de las estrictamente educativas. Además, la ponderación del nivel socioeconómico y cultural no sólo se efectúa cuando se comparan estudiantes, individualmente considerados, sino también cuando se evalúan los resultados logrados por colectivos más amplios, como pueden ser centros docentes o incluso sistemas educativos. Su propósito consiste en última instancia en asegurar la comparación justa entre las unidades evaluadas, evitando atribuir una eficacia mayor a aquellas unidades (centros educativos o distritos escolares) que simplemente se benefician de un alumnado más favorecido social y culturalmente. La realización de análisis que tienen en cuenta el denominado *valor añadido* no es sino uno de los ejemplos más claros y sofisticados de ese tipo de aproximación (Thomas, 1998).

Aunque en la actualidad se utilizan distintos modelos para calcular el valor añadido por los centros docentes, todos ellos combinan dos elementos fundamentales. El

primero consiste en la posición y el origen social de los estudiantes, que permite comparar los resultados obtenidos por un centro concreto en relación con los que cabría esperar estadísticamente en función de las características del alumnado que a él concurre. Es el modelo utilizado para calcular la eficacia de los liceos franceses que se utilizó en los años noventa (DEP, 1994). El segundo consiste en la comparación del resultado obtenido en dos momentos diferentes, lo que permite valorar el progreso realizado en una institución y tiempo concretos. Como puede apreciarse, los resultados más interesantes derivados de la aplicación de estas técnicas se consiguen acercando la escala de análisis al tamaño del centro docente, lo que no resulta factible en muchos estudios internacionales.

Los estudios internacionales del rendimiento educativo han tomado generalmente en consideración esta variable externa, aunque aplicando diversos procedimientos para analizar su efecto. La primera aproximación que se utilizó, todavía muy poco sofisticada, consistió en comparar los resultados logrados por los diversos países con algún indicador global de su grado de desarrollo económico, tal como el Producto Interior Bruto *per capita*. En este caso, los análisis se limitaron generalmente a yuxtaponer las tablas de los resultados obtenidos y del PIB nacional, o todo lo más a ajustar las puntuaciones de acuerdo con el PIB, como se aprecia en el ejemplo incluido en la Tabla 1.

TABLA 1

Resultados de los estudiantes, Producto Interior Bruto y puntuaciones ajustadas al PIB per capita

País	Puntuación PISA (2006)	Producto Interior Bruto (US\$) (2005)	Puntuaciones ajustadas al PIB per capita
Finlandia	563	30.923	558
Corea	522	22.277	534
Alemania	516	30.826	511
Reino Unido	515	32.890	506
Austria	511	34.409	499
Suecia	516	23.000	496
Dinamarca	496	34.090	485
Francia	495	30.352	491
EEUU	489	41.674	464
España	488	27.507	490
Italia	475	28.168	476
Portugal	474	20.043	490
Grecia	473	29.564	471
México	410	10.767	443

Fuente: OECD (2007, vol. 2: 45).

Otro procedimiento que también se ha utilizado habitualmente consiste en el cálculo de líneas de regresión entre los resultados y el PIB, derivadas de tablas como la mencionada, si bien conducen generalmente a análisis poco sofisticados. Así, por ejemplo, PISA 2006 introduce ese tipo de análisis, llegando a la conclusión de que el 28% de la varianza entre países se explica por el PIB *per capita* (OECD, 2007, vol. 1: 59).

Recientemente se ha realizado una aproximación más precisa al efecto del nivel socioeconómico y cultural, consistente en utilizar los datos individualizados de los alumnos evaluados, referidos a variables tales como el nivel de estudios y la ocupación profesional de los padres, los bienes materiales de que dispone la familia, el nivel de renta familiar, las expectativas académicas hacia los hijos o los recursos culturales disponibles en el hogar. La complejidad que supone obtener datos fiables de este tipo (por la pérdida de respuestas que suele producirse cuando se pregunta a las familias y por la imprecisión que implica muchas veces preguntar a los propios alumnos) ha sido el motivo de que hasta ahora se haya utilizado de forma limitada esta aproximación en los estudios internacionales⁴. El estudio PISA, que se dirige a jóvenes de 15 años, se ha animado a aplicar este tipo de instrumento, pese a las eventuales imprecisiones que pueda implicar, permitiendo análisis interesantes⁵.

TABLA 2

Relación entre los resultados de los estudiantes y su nivel socio-económico y cultural (PISA, medias brutas y ponderadas)

País	PISA 2000		PISA 2003		PISA 2006	
	MB	MP	MB	MP	MB	MP
Corea	525	534	542	549	522	522
España	493	504	485	504	488	499
EEUU	504	497	483	463	489	483
Francia	505	512	511	516	495	502
Italia	487	487	466	473	475	478
Portugal	470	488	466	508	474	492
Suecia	516	504	509	492	503	496
<i>Media OCDE</i>	<i>500</i>	<i>505</i>	<i>500</i>	<i>500</i>	<i>500</i>	<i>500</i>

Fuente: OECD (2001b: 308; 2004: 358 y 2007, vol. 1: 184).

4 Como se señalaba en la nota anterior, el estudio de Evaluación de Educación Primaria realizado por el INCE en 1995 sirvió para estudiar la aplicabilidad de los cuestionarios a los alumnos y a sus familias acerca de sus características socioeconómicas y culturales, arrojando unos resultados poco alentadores.

5 En este caso, la edad de los estudiantes evaluados ha resultado más conveniente para aceptar la fiabilidad de sus respuestas.

La Tabla 2 proporciona una información sintética de los resultados obtenidos en el Proyecto PISA, cuando se tiene en cuenta el nivel socioeconómico y cultural de los estudiantes evaluados, para los tres ciclos ya realizados y publicados. En cada una de las columnas de la izquierda correspondientes a cada uno de los tres años contemplados se presentan las medias brutas (MB) obtenidas por una selección de los países participantes, mientras que en las columnas de la derecha aparece su media ponderada (MP), obtenida mediante la aplicación del índice socioeconómico y cultural calculado a partir de los datos del alumnado.

Como puede apreciarse fácilmente, hay algunos países (España, Francia y Portugal) cuya media ponderada es en todos los casos superior a la bruta, lo que indica que sus resultados hubiesen sido superiores si los alumnos evaluados tuviesen una situación socioeconómica y cultural equivalente a la media de la OECD. Ese efecto es independiente de que sus resultados ponderados deban considerarse satisfactorios o no. Así, Francia se sitúa por encima de la media ponderada, mientras que España se coloca casi en ella y Portugal sigue quedando por debajo, aunque con una menor distancia que con respecto a la media bruta. El caso de Italia es interesante, pues su puntuación casi no cambia en 2000 al ponderarla, aunque queda algo más alejada de la media de la OECD al aumentar ligeramente esta última, mientras que mejora en 2003 y 2006. Una tendencia inversa presenta Corea, que mejora claramente su media al ponderarla en 2000 y 2003, pero ya no en 2006. Hay otros países, como Estados Unidos o Suecia, cuya puntuación ponderada es inferior a la bruta, como consecuencia de tener una población con un índice socioeconómico y cultural superior a la media de la OECD. En ambos casos la caída de posición es notable, pues Suecia pasa a situarse al nivel de España o algo por debajo, mientras que Estados Unidos cae claramente por debajo de la media ponderada de la OECD.

La lectura de esa tabla, aunque sea un simple ejemplo que no deba extrapolarse a cualquier otro ámbito, permite extraer varias conclusiones. Ante todo, resulta indudable que el nivel socioeconómico y cultural familiar influye en los resultados alcanzados por los alumnos. Pero también hay que reconocer que el efecto que produce la ponderación de esa variable es diferente para los diversos países, como no podría ser de otro modo. Así, para los países cuya población disfruta de unas condiciones socioeconómicas y culturales mejores que la media internacional, la ponderación produce un descenso de la puntuación, mientras que el efecto es el contrario para los países con una población más desfavorecida desde el punto de vista socioeconómico y cultural.

Por otra parte, no hay que creer que la ponderación de la situación socioeconómica y cultural de la población evaluada produzca necesariamente un efecto de igualación en las puntuaciones. El caso de Corea ilustra adecuadamente cómo en determinadas ocasiones la ponderación realizada eleva aún más la distancia con la media (que en este caso aumenta de 25 a 29 puntos por encima de la media de la OECD en 2000 y de 42 a 49 puntos en 2003, si bien permanece estable en 2006). El efecto concreto producido por este tipo de ponderación depende del nivel socioeconómico y

cultural de la población nacional, pero también del nivel de rendimiento logrado por el país. Por lo tanto, hay que concluir que los mejores resultados obtenidos por algunos países no se explican exclusivamente en función de esta variable extrínseca, debiendo incluir algunas otras en el análisis.

A la vista de estos análisis, cabe concluir que el nivel socioeconómico y cultural de los estudiantes y sus familias tiene un efecto importante en los resultados que aquellos logran, pero que en modo alguno es un efecto unívoco. Por lo tanto, hay que tenerlo en cuenta a la hora de explicar los resultados y de realizar interpretaciones de los mismos, pero no constituye una variable determinante, puesto que existen casos de países que, por arriba o por abajo, escapan a la lógica puramente socioeconómica.

— *Nivel de desarrollo educativo y cultural de la población*

Como se apuntaba más arriba, en los estudios nacionales que intentan valorar y explicar el rendimiento de los sistemas educativos se ha detectado un fenómeno interesante, que consiste en la relativa disociación del efecto de las condiciones socioeconómicas propiamente dichas de los alumnos, de un lado, y la riqueza de su entorno educativo y cultural, de otro (INCE, 1996). Dicho de otro modo, cuando se intenta explicar los resultados logrados por los estudiantes se aprecia que guardan una cierta relación con sus condiciones socioeconómicas, esto es, con los bienes materiales y la renta disponible de la familia, si bien la relación encontrada es bastante mayor si a esas variables se les agregan algunas otras sobre el acceso a bienes culturales, las aspiraciones y expectativas educativas familiares y las prácticas culturales, o sea, los elementos fundamentales de lo que Bourdieu y Passeron (1970) denominaron el *capital cultural* de la familia y los individuos.

Cabe suponer que ese fenómeno se debería también apreciar en los estudios internacionales, si bien hasta ahora no ha sido objeto de análisis detenido. En los trabajos promovidos por la IEA no se ha llevado a cabo dicho análisis hasta el momento actual, debido a las dificultades que se plantean para obtener la información necesaria⁶. El proyecto PISA, al obtener datos singularizados del nivel socioeconómico y cultural de los alumnos participantes, debería permitir hacer algún estudio de ese tipo, aunque todavía no se haya llevado a cabo de manera completa y rigurosa. Los únicos análisis disponibles combinan ambas dimensiones en un mismo índice, lo que impide apreciar su efecto diferencial.

En todo caso, parece lógico preguntarse por el efecto que pueda producir la existencia de condiciones educativas y culturales muy diferentes entre los países par-

⁶ Hay que tener en cuenta que muchos de los estudios promovidos por la IEA se dirigen a escolares de nueve o diez años de edad, lo que hace muy poco confiable la información de tipo socioeconómico que puedan proporcionar. Por otra parte, la distribución de cuestionarios a las familias plantea problemas importante, tanto logísticos como de tasa de respuesta, por lo que no se han utilizado habitualmente en estos estudios internacionales.

tipicantes en los estudios internacionales. Y además conviene preguntarse por el diferente efecto que pueda producir la riqueza socioeconómica y cultural de la familia propia, de un lado, y del entorno más amplio, de otro. Esto es, si el anterior conjunto de variables tenía una dimensión predominantemente individual y familiar, en este segundo grupo se abordan otras variables de dimensión colectiva y social. El presupuesto que subyace en tal distinción se basa en el fenómeno bien conocido del efecto que ejerce sobre el individuo el carácter más o menos estimulante del ambiente en que se desenvuelve.

En la Tabla 3 se han resumido algunos indicadores relativos al nivel educativo y cultural de la población de los países de la Unión Europea con quince miembros⁷. Un primer indicador seleccionado consiste en el porcentaje de la población comprendida entre los 15 y los 64 años de edad que ha finalizado al menos la educación secundaria superior (aunque bajo esa denominación común quepan realidades muy diversas). Otro indicador complementario consiste en el número de lectores habituales de periódicos por cada mil habitantes. Como tercer elemento de comparación, se ha incluido la puntuación bruta alcanzada por los países de la Unión Europea en el estudio PISA del año 2000, centrado en la comprensión lectora y que ha servido como criterio de ordenación⁸. Aunque los años de referencia de los tres indicadores no sean idénticos, no constituye un obstáculo insalvable para un análisis general como el que aquí se hace.

Lo primero que llama la atención es que los países europeos que obtienen en el proyecto PISA puntuaciones por encima de la media de la OCDE (500 puntos) tienen, en general, mayor porcentaje de población adulta con estudios superiores que los situados por debajo de esa media. Sin embargo, hay varias excepciones a esta regla general. Por una parte, Irlanda está incluida entre los países con menor proporción de población con estudios superiores (51%) y sin embargo obtiene una alta puntuación en PISA. Por otra parte, Alemania y Dinamarca son los países que tienen mayor proporción de población con estudios superiores (81% y 80%, respectivamente) y sin embargo quedan por debajo de la media. Es evidente que esa sencilla comparación no avala la existencia de una relación perfecta entre ambas variables. Pero si nos fijamos en los países del sur de Europa, comprobaremos que tanto Italia como España, Portugal y Grecia se encuentran entre los que menos población tienen con estudios superiores y también entre los que obtienen puntuaciones por debajo de la media, lo que nos obliga a no desechar rápidamente la existencia de algún tipo de relación.

7 Se ha preferido utilizar aquí los datos procedentes del estudio PISA, ya que en él han participado más países de la Unión Europea que en PIRLS. No obstante, se podrían y deberían realizar análisis comparativos de los resultados logrados en los distintos estudios internacionales, dadas las diferencias de enfoque, materia y población que implican, a las que se ha hecho mención con anterioridad.

8 Conviene hacer notar que, siendo PISA un estudio realizado por la OECD, las medias calculadas corresponden a los países miembros de tal organización. En consecuencia, los informes realizados no incluyen la media de los países de la Unión Europea, por lo que no se puede hacer ninguna comparación con la misma.

TABLA 3

Resultados de los estudiantes en comprensión lectora, población con estudios superiores y lectores de periódicos en los países de la Unión Europea-15

País	Puntuación PISA (2000)	Porcentaje de población 25-64 años con estudios superiores (1999)	Lectores de periódicos por 1000 habitantes (2000)
Finlandia	546	72	445
Irlanda	527	51	150
Reino Unido	523	62	319
Suecia	516	77	417
Austria	507	74	308
Bélgica	507	57	153
Francia	505	62	148
Dinamarca	497	80	277
España	493	35	106
Italia	487	42	104
Alemania	484	81	291
Grecia	474	50	64
Portugal	470	21	68
Luxemburgo	441	56	348
Holanda ⁹	–	–	278

Fuentes: OECD (2001^a: 45 y 2001b: 308), *El País* (18 abril 2002).

Por otra parte, al comparar los resultados de PISA con la lectura de periódicos vuelve a apreciarse un fenómeno similar. Ahora son tres los países incluidos en el grupo de menor proporción de lectores que quedan por encima de la media en PISA (Irlanda, Bélgica y Francia), mientras que dos de los países con mayor proporción de población lectora de periódicos quedan por debajo de la media (Luxemburgo y Alemania). Los países del sur de Europa vuelven a quedar por debajo de la media en ambos indicadores.

Aunque no es propósito de este trabajo descender a un análisis detenido y riguroso de esta posible relación, estudio que sin duda reclama atención y que deberá emprenderse en algún momento, una sencilla reflexión sobre los datos incluidos en la Tabla 3 apunta a la existencia de algún tipo de conexión entre el grado de desarrollo educativo y cultural de un país (medido aquí por dos indicadores simples) y los

⁹ No se incluyen datos de Holanda, dado que no aparecen en las dos publicaciones de la OECD que se han utilizado como fuente.

resultados logrados por sus estudiantes. Al igual que se observaba en el caso de la variable anterior, no podemos hablar de una relación perfecta, ni mucho menos de una determinación causal, pero hay indicios que obligan a analizar con más atención la conexión que pueda existir. En concreto, el análisis de la situación educativa de los países del sur de Europa, que han sufrido un retraso educativo histórico y que han evolucionado muy rápidamente en las últimas décadas, pero sin poder salvar fácilmente esa brecha centenaria que les separa de otros países europeos, merece un estudio más detenido, puesto que parece estar en relación con los resultados que alcanzan sus estudiantes. Desde ese punto de vista, valdría la pena hacer un estudio comparativo más cuidadoso con el caso de Irlanda, que podría arrojar información interesante.

— *Recursos destinados a la educación*

La tercera variable extrínseca aquí analizada se refiere a los recursos que cada país destina a la educación. Dichos recursos son de naturaleza muy variada, debiendo distinguirse al menos entre los humanos (profesorado, personal de administración y servicios educativos) y los económicos y materiales (infraestructura, equipamiento, presupuestos). Además, hay que tener en cuenta que esos recursos pueden proceder de muy diversas fuentes, tanto públicas como privadas, y de los distintos niveles de las administraciones públicas (estatal, provincial, regional, local). Puede incluso discutirse si esta variable debe considerarse extrínseca, puesto que entra dentro de la esfera propia de decisión de los gobiernos. El motivo de incluirla en este grupo consiste en que la determinación del presupuesto educativo es una tarea que no compete solamente a las autoridades educativas, sino que se inserta en un cuadro más complejo de asignación de prioridades y fondos presupuestarios para el conjunto de las políticas nacionales. Por lo tanto, no se trata de una decisión interna al propio sistema educativo, sino externa a él.

Las dificultades que plantea la comparación de los recursos realmente utilizados en las tareas educativas ha implicado que la mayor parte de los análisis hasta ahora realizados estén basados en el gasto que se realiza en educación (incluyendo a veces la inversión propiamente dicha), medido en términos monetarios, siendo mucho menos frecuente otro tipo de comparación. Además, hay que señalar que los estudios internacionales de evaluación del rendimiento raramente incluyen análisis precisos acerca de la relación entre los resultados educativos y los recursos monetarios utilizados. Lo más que hacen, como en el caso de TIMSS o del proyecto PISA, es yuxtaponer ambas informaciones, pero sin analizar detenidamente su relación. Por lo tanto, al igual que sucedía en el apartado anterior, no es fácil llegar a conclusiones definitivas sobre la conexión que pueda existir entre ambas variables.

A pesar de esa dificultad, también en este caso he intentado analizar el efecto que ejerce la variable que aquí se considera. En la Tabla 4 he recogido algunos datos

correspondientes a varios de los países participantes en el proyecto PISA¹⁰. La selección de países se ha efectuado de acuerdo con la relevancia que dichos casos tienen en el contexto de este trabajo y con su carácter ejemplificador. La primera columna incluye los resultados brutos obtenidos en PISA 2000 que, al igual que en otras tablas precedentes, han servido para ordenar los países. La segunda columna recoge el Producto Interior Bruto *per capita* de dichos países, mientras que la tercera incluye el gasto acumulado en instituciones educativas por estudiante. Estas dos últimas cifras se expresan en dólares USA, utilizando un algoritmo de paridad del poder adquisitivo (PPP), a fin de hacer las cifras comparables. La última columna incluye el cociente resultante de dividir las dos anteriores, representando el esfuerzo económico efectuado por estudiante en relación con el Producto Interior Bruto nacional.

TABLA 4

Resultados de los estudiantes, Producto Interior Bruto y gasto acumulado en instituciones educativas por estudiante

País	Puntuación PISA (2000)	Producto Interior Bruto (US\$) (1999)	Gasto acumulado por estudiante (US\$) (1998)	Cociente Gasto acumulado / PIB
Finlandia	546	22.800	45.363	1,99
Corea	525	15.900	30.844	1,94
Reino Unido	523	22.300	42.793	1,92
Suecia	516	23.000	53.386	2,32
Austria	507	24.600	71.387	2,90
Francia	505	21.900	50.481	2,31
EEUU	504	33.900	67.313	1,99
Dinamarca	497	26.300	65.794	2,50
España	493	18.100	36.699	2,03
Italia	487	21.800	60.824	2,79
Alemania	484	23.600	41.978	1,78
Grecia	474	14.800	27.356	1,85
Portugal	470	16.500	36.521	2,21
México	422	8.100	11.239	1,39

Fuente: OECD (2001b: 264).

¹⁰ En este caso he preferido utilizar también los datos de PISA en vez de los de TIMSS, por su mayor cobertura europea (como ocurría en la Tabla 3) y por proporcionar en el mismo informe los datos sobre gasto acumulado por alumno, lo que implica la aplicación de una metodología, si no idéntica, al menos coherente entre ambas series de datos.

Como puede apreciarse en la tabla, los resultados de los países no siguen el mismo orden que el gasto acumulado por estudiante. Así, Corea se sitúa en la parte alta de la tabla, pese a que su gasto por estudiante es de los más bajos de la OCDE, exactamente al contrario que le sucede a Italia. Además, hay países con un nivel similar de rendimiento en PISA 2000, a los que corresponden niveles bastante dispares de gasto. Es lo que sucede al comparar Francia con Estados Unidos, Corea con el Reino Unido o Grecia con Portugal. A la inversa, hay países con niveles comparables de gasto acumulado por estudiante pero con niveles dispares de resultados, como el caso del Reino Unido y Alemania o España y Portugal.

Tampoco el cociente que se obtiene al dividir el gasto acumulado por estudiante entre el PIB sigue un orden igual al de los resultados. Aunque México ocupa la última posición en ambos casos, Italia está muy por debajo de lo que le correspondería en el caso de existir una correlación perfecta, y también estarían fuera de su lugar países tan dispares como Corea, el Reino Unido y Finlandia (todos ellos por encima de su posición teórica), así como Portugal y Dinamarca (en este caso por debajo).

Por lo tanto, también en el caso de esta tercera variable extrínseca pueden hacerse algunas consideraciones similares a las que se formulaban con las dos anteriores. Nos encontramos en una situación en la que podemos afirmar que todo apunta a la existencia de una cierta relación entre los recursos dedicados a la educación y el rendimiento educativo, pero la relación dista mucho de ser lineal ni perfecta. Sin duda, es una de las variables que deben ser tenidas en cuenta a la hora de valorar y explicar los resultados, pero no es la única. Como se ha afirmado muchas veces, no se trata solamente de dedicar más recursos a la educación, sino de emplearlos bien. Y en esta última parte de la afirmación intervienen otras variables, muchas de ellas de carácter intrínseco, que han de tenerse en cuenta.

Antes de entrar en el análisis de las variables intrínsecas, habría que señalar que las reflexiones realizadas en este apartado, concordantes por otra parte con una orientación ya tradicional en la investigación educativa, apuntan a la existencia de relaciones múltiples y cruzadas entre todas estas variables extrínsecas. Por lo tanto, aun cuando pueda interesar en ocasiones el estudio singular de algunas de ellas, no es posible explicar los resultados logrados en los estudios internacionales por medio de una sola variable. Los análisis deben respetar la complejidad de los fenómenos estudiados e integrar distintos factores en cada explicación.

3. EN BUSCA DE MÁS EXPLICACIONES: VARIABLES INTRÍNSECAS DEL RENDIMIENTO

Las variables extrínsecas ayudan a explicar los resultados conseguidos por los estudiantes en las pruebas de rendimiento escolar, pero no dan cuenta total

de los mismos. Por ese motivo hay que explorar otro tipo de factores, relativos a las características y al funcionamiento del sistema educativo, que tienen además especial interés cuando se pretende llevar a cabo acciones de mejora, ya que caen plenamente dentro de la esfera de actuación de las autoridades educativas, de los centros escolares y del profesorado. Son las aquí denominadas *variables intrínsecas*.

Las variables intrínsecas que analizan los estudios actuales de evaluación del rendimiento son muy diversas. Muchos de los estudios promovidos por IEA, OECD y UNESCO recogen información acerca de variables tales como la formación y el desempeño de los profesores, el grado de autonomía de la escuela, el tiempo dedicado a la enseñanza y al estudio personal, la organización del currículo, el tipo de refuerzo prestado a los alumnos y otras similares, dimensiones todas ellas de la vida educativa que caen dentro de esta categoría.

Entre este tipo de variables he seleccionado aquí tres grupos. En primer lugar, analizaré los efectos producidos por algunos factores relacionados con la organización general del sistema educativo. En segundo lugar, pasaré revista al efecto de los denominados *factores de centro escolar*, que se refieren a la organización, el funcionamiento y el clima o ambiente escolar, considerados en términos generales. En tercer lugar, analizaré el efecto de los denominados *factores de aula*, que guardan relación con las condiciones en que se desarrollan los procesos de enseñanza y aprendizaje en el interior del aula, en una situación de contacto directo entre el profesor y los alumnos.

— *Organización del sistema educativo*

Las variables intrínsecas de mayor nivel de generalidad son las que se refieren a la organización del sistema educativo. Se trata de variables que afectan al conjunto de los centros y de los agentes escolares, puesto que tienen que ver con las normas fundamentales y las condiciones generales en que aquellos deben desenvolverse. Este tipo de variables son en realidad características relativas a la estructura del propio sistema educativo y a sus reglas básicas de funcionamiento.

De acuerdo con los supuestos de la concepción del *mundo como laboratorio educativo* que se comentaba al comienzo del trabajo, los estudios internacionales deberían contribuir a proporcionar explicaciones ligadas precisamente a este tipo de variables, ya que son las que distinguen más claramente a unos sistemas educativos de otros. Sin embargo, hay que reconocer que los estudios realizados rara vez son concluyentes y casi siempre proporcionan una información que requiere interpretaciones complementarias. Hasta ahora son pocas las conclusiones definitivas que han aportado los estudios internacionales sobre este tipo de variables. Sin embargo, se pueden señalar algunas, aunque haya que tomarlas con cierta precaución.

Entre los ejemplos más relevantes de este tipo de análisis, hay que acudir una vez más a PISA. Uno de los fenómenos que más llamó la atención tras la publicación del informe de resultados correspondiente al año 2000, y que más debate generó, consistió en el pobre rendimiento educativo de los jóvenes alemanes de quince años. Como se ha visto en las tablas insertas en el trabajo, su puntuación está claramente por debajo de la media de la OECD y también por debajo de lo que cabría esperar de un país con un alto nivel de desarrollo. Ese dato provocó alarma y generó una viva discusión (Kotthoff y Pereyra, 2009). Lo que no llamó tanto la atención es que esa circunstancia no era exclusiva de Alemania, sino que también se apreciaba, en mayor o menor grado, en otros países con sistemas educativos semejantes. La característica fundamental de ese tipo de sistemas consiste en una *diferenciación institucional temprana*. Dicho de otro modo, los adolescentes de esos países son separados muy tempranamente (entre los diez y los doce años de edad) en itinerarios formativos muy diferenciados, que se corresponden con otros tantos modelos institucionales escolares. Aunque el sistema prevé la posibilidad de trasvases entre unos y otros itinerarios, ese paso se produce en la práctica en muy pocos casos.

A partir de estos datos, los responsables del estudio PISA llegaron a una triple conclusión que merece la pena destacar (OECD, 2001b). En primer lugar, las diferencias que existen entre los diversos tipos de escuelas en esos países son generalmente mayores que las que se observan en otros lugares. Si en la mayor parte de los países participantes en el estudio la varianza intra-escuelas era mayor que la varianza entre escuelas, en estos países sucedía exactamente lo contrario. En segundo lugar, la variación de resultados que se aprecia entre los estudiantes evaluados también suele ser mayor que en otros países, con muy pocas excepciones. En tercer lugar, ninguno de esos países se encuentra entre los de mayor rendimiento promedio. O sea, el estudio PISA parece concluir que la discriminación institucional temprana (la que tiene lugar antes de los doce años de edad) produce desigualdades importantes en los países que la aplican, sin llegar por ello a obtener unos resultados promedios satisfactorios. Los mejores estudiantes de esos sistemas educativos están entre los mejores del mundo, pero también los peores se encuentran entre los peores, provocando una situación global poco positiva y desequilibrada. Cabría inferir que este modelo de organización del sistema educativo parece poco aconsejable para las circunstancias actuales, tanto por la escasa equidad que logra como por su bajo rendimiento.

Otro ejemplo concreto del análisis de ciertas variables que parecen incidir en los resultados de los estudiantes se encuentra en las conclusiones sobre la autonomía escolar que aporta el estudio PISA. A partir de la información recogida en dicho estudio, se ha elaborado la Tabla 5, que sintetiza la relación existente entre los resultados que logran los alumnos y el porcentaje de directores escolares que afirman tener al menos alguna autonomía en una serie de áreas de actuación educativa.

TABLA 5

*Autonomía escolar y resultados educativos
(relación entre los resultados de los alumnos y el porcentaje de directores
que tienen alguna responsabilidad en los siguientes aspectos de la gestión escolar)*

Decisión sobre los cursos ofrecidos	0,51
Determinación de asignaciones presupuestarias	0,37
Elección de manuales escolares	0,30
Determinación de los contenidos de los cursos	0,25
Establecimiento de la política disciplinaria	0,21
Establecimiento de la política de evaluación de los alumnos	0,20
Nombramiento de profesores	0,16
Despido de profesores	0,10

Fuente: OECD (2001b: 175).

Lo que la tabla manifiesta es que no se puede afirmar taxativamente que una mayor o menor autonomía escolar produzca globalmente mejores o peores resultados de los estudiantes. Como puede apreciarse a partir de esos datos, la autonomía escolar es una realidad que abarca diversas dimensiones y no todas ellas tienen la misma incidencia sobre el rendimiento escolar. En términos generales, podría decirse que la autonomía en los aspectos relativos a la organización curricular, la vida y el ambiente del centro escolar y la gestión de sus recursos presupuestarios parece ejercer un mayor efecto sobre el rendimiento que la autonomía en lo relativo a la gestión del personal docente. Así, cabría concluir que ciertas dimensiones de la autonomía escolar merecen un especial apoyo por parte de las autoridades educativas, mientras que hay otros aspectos en que la acción no es tan prioritaria.

— *Procesos de centro escolar*

Este último tipo de reflexiones conducen directamente al segundo grupo de variables intrínsecas, las relacionadas con la vida de los centros escolares. Son los comúnmente denominados *procesos de centro*. Los procesos de centro y los de aula, que serán objeto de comentario en el apartado siguiente, constituyen en conjunto lo que habitualmente se conoce como *procesos escolares*. Esta dimensión constituye un nivel intermedio entre el contexto socioeconómico y cultural y los resultados educativos. O sea, la acción desarrollada por los profesores y por otros miembros de la comunidad escolar, mediante el desarrollo de diversos procesos y actuaciones, es la que explica que se logren unos determinados resultados, a partir de unas determinadas condiciones de partida. De ahí deriva su importancia y el interés que el tema despierta en los investigadores (Marchesi y Martín, 1998).

La importancia de este tipo de procesos es indudable, aunque haya que situarla en su justo término. Frente a la tendencia inaugurada por la publicación del célebre Informe Coleman (Coleman *et al.*, 1966) y por la sociología crítica que insiste en la *reproducción* (Bourdieu y Passeron, 1970), la investigación educativa reciente ha puesto de manifiesto que no todos los resultados de la educación pueden explicarse exclusivamente en función de las fuerzas reproductoras del contexto, aunque su presencia y su impacto sean innegables. Así, un estudio de hace unos años en que se revisaron veinticuatro investigaciones y se realizaron varios metaanálisis llegaba a la conclusión de que entre un 12% y un 15% de la varianza en los resultados que alcanzan los alumnos es debida a la acción del centro escolar (Teddlie, Reynolds y Sammons, 2000). Se trata de un porcentaje que, sin ser demasiado elevado, no resulta despreciable y que está en línea con lo que otros investigadores han concluido. Podríamos decir que ese es el espacio propio de actuación de los centros educativos y del profesorado y que, por tanto, tiene notable importancia desde el punto de vista pedagógico.

Los estudios de evaluación del rendimiento de ámbito nacional han concedido importancia a los procesos de centro y de aula. Los estudios internacionales también han incluido este tipo de procesos entre sus variables de análisis, aunque con unos enfoques y una formulación ligeramente diferente de uno a otro caso. Además, algunos proyectos concretos de evaluación institucional, como el desarrollado en España por el Instituto IDEA, han prestado notable atención a su estudio, en la convicción de que constituyen el ámbito privilegiado de cualquier plan de actuación para la mejora de la educación (Marchesi y Martín, 2002).

Hablando concretamente de los procesos de centro escolar, he revisado las variables de proceso contempladas por los estudios desarrollados en las últimas décadas por la IEA, OECD y UNESCO antes mencionados, con el fin de identificar las que parecen guardar mayor relación con el rendimiento. Entre todas las variables incluidas en esos estudios, las que parecen gozar de un mayor consenso son las siguientes:

- a) Liderazgo pedagógico, entendido como la existencia de una dirección con autoridad pedagógica reconocida, capaz de promover un buen funcionamiento del centro escolar, facilitadora del logro de un buen clima escolar y eficaz en su funcionamiento cotidiano.
- b) Coordinación pedagógica y curricular entre los profesores y las etapas, con el fin de asegurar una coherencia interna en la enseñanza proporcionada y en los medios puestos en juego para lograr el aprendizaje.
- c) Implicación familiar en la escuela, como señal de la participación de los padres en la educación de sus hijos y de un estilo de actuación capaz de asegurar la coherencia educativa entre la escuela y el hogar.
- d) Buen ambiente o *clima* escolar, entendido como la existencia de unas buenas relaciones entre los distintos miembros de la comunidad escolar (estudiantes, profesores, familias), facilitadoras de la tarea educativa.

- e) Formación y estabilidad del profesorado, ya que existe la certeza de que la actuación de los docentes constituye un elemento capital para el logro de una enseñanza de calidad, y que esa actuación no puede ser positiva si no se dan unas condiciones favorables, tanto en lo que se refiere a la capacitación de los profesores, como a su seguridad laboral, una de cuyas variables fundamentales es la estabilidad (Barber y Mourshed, 2007).
- f) Tamaño adecuado de la escuela, que no siempre resulta fácil de determinar, pero que tiene relación con la existencia de una ambiente diversificado y rico (que exige unas dimensiones mínimas) pero no masificado (que impone unos límites superiores).

El modo en que los distintos estudios internacionales formulan y exploran el efecto de estas variables difiere de unos a otros casos, pero existe una notable coincidencia en la elección de las áreas abarcadas. Aún a falta de realizar un análisis comparado detenido de esas formulaciones y aproximaciones, cuestión que tiene indudable interés pero aún no se ha llevado a cabo, se aprecia la existencia de unas áreas prioritarias de análisis en dichos estudios. No obstante, hay que señalar que la dispersión de aproximaciones utilizadas por las diferentes organizaciones que los promueven, la dificultad que plantea la definición común de variables como las mencionadas para sistemas educativos tan distintos entre sí y la limitada consistencia de los análisis realizados, obligan a seguir trabajando en este campo, como han hecho tanto la IEA como la OECD en sus estudios respectivos¹¹.

— *Procesos de aula*

Las variables referidas a los procesos de centro se complementan estrechamente con las relativas a los denominados *procesos de aula*. Si las variables anteriores configuran un campo complejo y multidimensional, en el que no resulta sencillo encontrar explicaciones unívocas y obliga a combinarlas en análisis multifactoriales, esa complejidad se acentúa incluso más cuando descendemos al plano de las prácticas educativas en el aula. Como afirma Elena Martín, lo que sucede en una clase se caracteriza por la multidimensionalidad, la simultaneidad, la inmediatez, la impredecibilidad, la singularidad y el carácter valorativo de las actividades (Martín, 2002). Por ese motivo, todos los estudios que intentan valorar el efecto de este tipo de variables en el rendimiento de los alumnos tienen que reducir su foco de atención y centrarse en algunas variables concretas.

Si eso sucede en las investigaciones que adoptan el aula como objeto de estudio, aún con mayor motivo se ven obligados a hacerlo los estudios internacionales, dada la

¹¹ Un ejemplo destacable de este empeño es el estudio denominado TIMSS-Video, que desarrolló la IEA en conexión con la aplicación de TIMSS y que estaba basado en el análisis de clases de matemáticas y ciencias grabadas en video en varios países, con la intención de analizar en profundidad los diversos procesos aplicados para la enseñanza de esas materias, sin limitarse a procesar las respuestas que los propios profesores dan a través de cuestionarios (Hiebert *et al.*, 2003). La información obtenida permitió completar y contrastar los datos aportados por los propios profesores, enriqueciendo así el análisis.

distancia que mantienen con las aulas y la diversidad de estilos de configuración y funcionamiento de éstas que existen en el plano internacional. Por esa razón, dichos estudios se centran en un número limitado de variables, seleccionando aquellas que la investigación educativa reconoce como más potentes a la hora de explicar los resultados.

Entre las variables relativas a los procesos de aula que se han tomado en consideración en los estudios internacionales realizados durante las dos últimas décadas y que han arrojado unos resultados más prometedores, he identificado las siguientes:

- a) Tiempo dedicado a la enseñanza, entendido tanto en el sentido del tiempo efectivamente dedicado a las tareas docentes en la escuela (excluyendo, por ejemplo, interrupciones no deseadas, que en algún caso llegan a ser considerables), como el tiempo dedicado a tareas escolares en el hogar.
- b) Estructuración de la enseñanza, entendida como una organización adecuada y cognitivamente estimulante de las tareas de aprendizaje por parte del docente.
- c) Oportunidad de aprender, entendida en el sentido en que la IEA desarrolló este concepto hace tiempo, esto es, como la exposición real a la enseñanza (en el sentido del currículo efectivamente impartido), que proporciona al estudiante la oportunidad de desarrollar su aprendizaje.
- d) Altas expectativas de rendimiento por parte de los profesores, factor que la investigación educativa ha demostrado cumplidamente que va ligado a unos mejores resultados de los estudiantes.
- e) Evaluación continua y control del progreso de los alumnos, en el sentido de ofrecer una realimentación permanente del proceso de enseñanza y aprendizaje y de ayudar a los alumnos a valorar por sí mismos su progreso.
- f) Refuerzo positivo al aprendizaje, entendido de diversos modos, no necesariamente en su concepción conductista, y que tiene que ver con la introducción de estímulos para continuar en la tarea y para esforzarse en conseguir mejores resultados.
- g) Proporción adecuada de alumnos por profesor que, al igual que sucedía con el tamaño de la escuela, no es una función lineal, sino que tiene un mínimo y un máximo convenientes, por debajo o por encima de los cuales la tarea docente y el aprendizaje resultan menos estimulantes o más difíciles.

El problema que plantean todas estas variables, tanto las referidas a los procesos de centro como a los procesos de aula, radica en la dificultad que implica su investigación empírica. Los docentes y los estudiosos de la educación saben que se trata de variables muy influyentes en el desarrollo de las actividades escolares y, en conse-

cuencia, en los resultados logrados por los alumnos. Pero dada su relación multidimensional, no resulta sencillo medir ni valorar su efecto concreto.

Una de las líneas de investigación que ha pretendido estimar el efecto de este tipo de variables sobre el rendimiento académico es la que se conoce como estudio de la *eficacia escolar*. Se trata de una línea de trabajo con una larga tradición, pero que se renovó en los años noventa (Scheerens, 1992). Entre los principales autores que han contribuido a renovarla merece la pena citar al holandés Jaap Scheerens. En un trabajo publicado en 1996, uno de los pocos existentes que han intentado realizar un análisis global de los factores de eficacia escolar, se preguntaba por la aportación que este movimiento puede realizar para la gestión y la mejora de la educación (Scheerens, 1996). Lo cierto es que sus conclusiones no resultaban excesivamente esperanzadoras. Tras realizar diversos análisis empíricos a partir de investigaciones propias y de otros autores, llegaba a la conclusión de que muchas de las variables acerca de cuyo efecto existe un consenso bastante generalizado no disfrutaban de una base empírica suficiente, como recoge la Tabla 6.

TABLA 6

Grado en el cual las variables intrínsecas se han visto confirmadas por la investigación empírica

Característica	Confirmación múltiple	Base empírica razonable	Confirmación dudosa	Hipotética
Enseñanza estructurada	X			
Tiempo de aprendizaje	X			
Oportunidad de aprender		X		
Orientación hacia el logro		X		
Altas expectativas		X		
Liderazgo pedagógico			X	
Habilidad evaluadora			X	
Clima escolar			X	
Reclutamiento de personal				X
Condiciones estructurales y organizativas			X	
Condiciones físicas y materiales		X		
Características del contexto			X	
Estímulos externos hacia la eficacia escolar				X
Implicación familiar		X		

Fuente: Scheerens (1996: 111).

La lectura correcta de esta tabla no implica que dichas variables deban ser despreciadas ni que su efecto sea irrelevante. Sabemos que muchas de ellas desempeñan un papel importante en la actividad escolar. El problema está en que no resulta fácil cuantificarlas ni medirlas. Y si esa dificultad surge en estudios de pequeña escala, aún es mayor en los estudios internacionales, que adoptan una perspectiva muy amplia. Por ese motivo, los resultados que hasta ahora ha proporcionado en este campo la investigación internacional del rendimiento son limitados. Aunque todas las organizaciones que trabajan en este campo están desarrollando modelos finos y sofisticados para valorar el impacto de las variables intrínsecas sobre el rendimiento, todavía no contamos con aproximaciones unánimemente aceptadas.

4. CONCLUSIONES

Como se ha analizado en estas páginas, los estudios internacionales de evaluación del rendimiento educativo surgieron con la intención primordial de permitir una comparación rigurosa de los resultados logrados por los estudiantes de los diversos sistemas educativos. No cabe ninguna duda de que ese objetivo se ha saldado con éxito. En efecto, este tipo de estudios, que ya cuentan con una considerable tradición de casi cincuenta años, permiten comparar la situación de un sistema educativo determinado con otros que pueden servirle de referencia, dando la posibilidad de identificar sus principales puntos fuertes y débiles, así como de estimar el progreso que se produce a lo largo del tiempo. Hoy en día, la información proporcionada por proyectos como TIMSS, PIRLS o PISA es objeto de estudio especializado, de debate público, de análisis académico y de valoración política. La comparación de los sistemas educativos se ha enriquecido notablemente con la información suministrada por los sistemas internacionales de indicadores y por los estudios de evaluación de gran escala.

Pero dichos estudios también pretendieron desde el comienzo contribuir a explicar los resultados conseguidos, por medio del análisis de los denominados *factores asociados*. Es en este aspecto en el que me he detenido en estas páginas y sobre el que conviene plantear algunas reflexiones finales.

Quizás el principal problema que he abordado aquí no se refiera sobre todo a qué explicaciones han proporcionado hasta ahora los estudios internacionales del rendimiento, sino a qué podemos legítimamente esperar de ellos. Hasta ahora, podemos afirmar que tales estudios han arrojado luz sobre algunos de los factores analizados. Así, hoy poseemos bastantes más certezas que hace unos años acerca de la relación entre los resultados educativos y el nivel socioeconómico y cultural de los estudiantes, las familias y el entorno en que se desenvuelven. No es que sea una novedad absoluta, pero los análisis comparativos internacionales han ayudado a clarificar la cuestión. También hemos encontrado alguna confirmación estadística, basada en el manejo de grandes bases de datos, del efecto producido por características estructurales del sistema educativo, tales como la diferenciación institucional temprana o las diversas dimensiones de la autonomía escolar. Y lo mismo podemos decir de

los principales efectos sobre el rendimiento de los factores contemplados por el movimiento de la eficacia escolar. Pero, sin embargo, la impresión que nos queda al final del análisis es de cierta insatisfacción, pues tenemos más preguntas y dudas que respuestas y certezas. Podemos decir que hemos encontrado algunas nuevas explicaciones, pero aún no podemos explicar convincentemente y con rigor el rendimiento educativo.

Desde este punto de vista, se puede decir que la concepción del *mundo como laboratorio educativo* que alumbró la IEA en los años sesenta fue una idea feliz, pero más como inspiración que como resultado. No es justo considerar que se trata de una quimera, pero sí de un planteamiento tendencial, hacia el que nos debemos orientar y que quizás no alcancemos nunca.

No cabe duda de que los estudios internacionales ofrecen algunas posibilidades interesantes, como las arriba señaladas, aunque a condición de reconocer sus limitaciones. En mi opinión, su potencia explicativa se ve limitada sobre todo por la perspectiva macroscópica que adoptan, que les hace perder detalles importantes, por la mirada poco contextualizada que necesariamente deben arrojar sobre los sistemas educativos y por la estandarización que imponen sus métodos y técnicas de análisis. Cuando se trata de estudiar fenómenos singulares, de pequeña escala o que deben estar correctamente contextualizados, la opción de los estudios internacionales puede no ser la más adecuada.

Aunque aquí no se han analizado otras iniciativas de evaluación, como los estudios nacionales de diagnóstico o la evaluación institucional, no cabe duda de que constituyen un complemento muy valioso, incluso indispensable, para comprender y valorar más profundamente las razones y los factores que explican los resultados alcanzados. Por más valiosos e interesantes que sean los estudios internacionales, que sin duda lo son, no bastan para comprender el funcionamiento profundo del sistema educativo ni para adoptar todas las decisiones que redunden en la mejora de la educación. De ahí deriva el interés que tiene la comprensión de la evaluación como una tarea integrada por acciones de diversa naturaleza y con diverso enfoque y perspectiva, idea cada vez más compartida.

De esta serie de reflexiones se deduce la necesidad que tienen los sistemas educativos de adoptar instrumentos adecuados para llevar a cabo la gestión de sus recursos y el seguimiento de sus resultados (Tiana, 2003). Conocer los resultados y su explicación se ha convertido cada vez más en una exigencia para la mejora. No es casualidad que en este contexto se hayan desarrollado nuevos conceptos, como el de *pilotage* o *steering* que, en castellano, quizás podría traducirse por *conducción*. La denominación fue introducida por Gilbert de Landsheere, para quien podría definirse como “la toma de decisiones al nivel macroscópico (es decir, de un sistema o subsistema educativo) o microscópico (es decir, de un establecimiento o incluso una clase), sobre la base de constataciones tan objetivas como sea posible, relativas al estado, al funcionamiento o a los productos del sistema” (De Landsheere, 1994:8). La evaluación

es un instrumento muy adecuado y valioso para proporcionar ese tipo de información objetiva y confiable (Michel, 1996).

Pero, si se concibe la conducción como una tarea que se desarrolla a diversos niveles, los estudios internacionales no resultan suficientes para realizarla de forma completa. Los estudios internacionales de evaluación del rendimiento no son sino uno de los instrumentos disponibles, pero en modo alguno el único. Los países pueden obtener de ellos información relevante sobre su situación educativa, pero necesitan combinarla con otro tipo de datos y de análisis. Ahí es donde se encuentra el interés y las posibilidades reales que ofrecen los estudios internacionales, y por su contribución a esa tarea habrá que valorarlos.

REFERENCIAS BIBLIOGRÁFICAS Y DOCUMENTALES

- Álvaro, M. *et al.* (1990), *Hacia un modelo causal del rendimiento académico*. Madrid: Ministerio de Educación y Ciencia – Centro de Investigación, Documentación y Evaluación.
- Barber, M. y Mourshed, M. (2007), *How the world's best-performing school systems come out to the top*. Londres: McKinsey & Co.
- Bourdieu, P. y Passeron, J.C. (1970), *La reproducción: éléments pour une théorie du système d'enseignement*. Paris: Editions de Minuit.
- CERI (1994), *Making Education Count. Developing and Using International Indicators*. Paris: CERI – OECD.
- Coleman, J.S. *et al.* (1966), *Equality of Educational Opportunity*. Washington D.C.: Office of Education.
- De Landsheere, G. (1994), *Le pilotage des systèmes d'éducation*. Bruselas: De Boeck.
- Degenhart, R.E., ed. (1990), *Thirty years of international research. An annotated bibliography of IEA publications (1960-1990)*. The Hague: IEA.
- DEP (1994), *Trois indicateurs de performance des lycées, Baccalauréat 1993, Résultats lycée par lycée*. Paris: Ministère de l'Éducation nationale – Direction de l'Évaluation et de la Prospective (2 tomos).
- Hiebert, J., *et al.* (2003), *Teaching Mathematics in Seven Countries: Results From the TIMSS 1999 Video Study (NCES 2003-013)*. Washington DC: United States Department of Education - National Center for Education Statistics.
- INCE (1996), *Evaluación de la Educación Primaria. Informe preliminar*. Madrid: Instituto Nacional de Calidad y Evaluación (documento policopiado).
- Kotthoff, H-G. y Pereyra, M.A. (2009), “La experiencia del PISA en Alemania: Recepción, reformas recientes y reflexiones sobre un sistema educativo en cambio”, *Profesorado. Revista de currículum y formación del profesorado*, vol. 13, 2: 1-24.
- Marchesi, A. y Martín, E. (1998), *Calidad de la enseñanza en tiempos de cambio*. Madrid: Alianza Editorial.
- Marchesi, A. y Martín, E., comps. (2002), *Evaluación de la educación secundaria. Fotografía de una etapa polémica*. Madrid: Fundación Santa María – SM.

- Martín, E. (2002), “Los procesos de aula y su influencia en la calidad de la enseñanza”, en A. Marchesi y E. Martín, comps., *Evaluación de la educación secundaria. Fotografía de una etapa polémica*. Madrid: Fundación Santa María – SM.
- Michel, A. (1996), “La conducción de un sistema complejo: la Educación Nacional”, *Revista Iberoamericana de Educación*, n.º 10: 13-36.
- OECD (1992), *High-Quality Education and Training for All*. Paris: OECD.
- (2001a), *Education at a Glance. OECD Indicators*. Paris: OECD.
- (2001b), *Knowledge and Skills for Life. First Results from PISA 2000*. Paris: OECD.
- (2004), *Learning for Tomorrow’s World. First results from PISA 2003*. Paris: OECD.
- (2007), *PISA 2006. Science Competencies for Tomorrow’s World*. Paris: OECD, 2 vols.
- OECD & Statistics Canada (1995), *Literacy, Economy and Society. Results of the first International Adult Literacy Survey*. Paris: OECD – Ministry of Industry of Canada.
- (2000), *Literacy in the information age: final report of the international adult literacy survey*. Paris y Ottawa: OECD – Statistics Canada.
- OREALC (1998), *Primer estudio internacional comparativo sobre lenguaje, matemática y factores asociados en tercero y cuarto grado*. Santiago de Chile: OREALC – UNESCO.
- (2008), *Los aprendizajes de los estudiantes de América Latina y el Caribe. Primer reporte de los resultados del Segundo Estudio Regional Comparativo y Explicativo*. Santiago de Chile: OREALC.
- Papadopoulos, G.S. (1994), *L’OCDE face à l’éducation, 1960-1990*. Paris: OCDE.
- Papanastasiou, C., ed. (2004), *Proceedings of the International Research Conference 2004*. Nicosia: University of Cyprus – IEA, 3 vols.
- Purves, A. (1993), “The World as an Educational Laboratory”, en W.A. Hayes, ed., *IEA Activities, Institutions and People. IEA Guidebook 1993-1994*. The Hague: IEA.
- Scheerens, J. (1992), *Effective Schooling: Research, Theory and Practice*. Londres: Cassell.
- (1996), “Can the School Effectiveness Knowledge Base Guide School Management?”, en *II International Conference on School Management. Participative Management and School Evaluation*. Bilbao: Universidad de Deusto.
- Teddlie, Ch., Reynolds, D. y Sammons, P. (2000), “The Methodology and Scientific Properties of School Effectiveness Research”, en D. Reynolds y Ch. Teddlie, eds., *The International Handbook of School Effectiveness Research*. Londres: The Falmer Press.
- Thomas, S. (1998), “Value-added measures of school effectiveness in the United Kingdom”, *Prospects*, vol. XXVIII, 1: 91-108.
- Tiana, A. (2001), “Le monde comme laboratoire éducatif”, *Politiques d’éducation et de formation. Analyses et comparaisons internationales*, vol. 2001/3: 47-57.
- (2002), “El contexto sociocultural en la evaluación de los centros educativos”, en A. Marchesi y E. Martín, comps., *Evaluación de la educación secundaria. Fotografía de una etapa polémica*. Madrid: Fundación Santa María – SM.
- (2003), “Pilotage par les résultats et amélioration de l’éducation: deux réalités indissociables?”, *Administration et Éducation*, 98: 19-29.

- (2006), “Assessing Quality in Education: Concepts. Models and Instruments”, en Dobbelstein, P. y Neidhart, T., eds., *Schools for Quality – What data-based Approaches Can Contribute*. Sint-Katelijne-Waver: Consortium of Institutions for Development and Research in Education in Europe (CIDREE).