



VOL. 13, Nº 2 (2009)

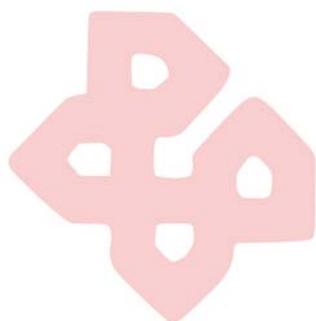
ISSN 1138-414X

Fecha de recepción 14/02/2009

Fecha de aceptación 23/07/2009

EL VALOR INFORMATIVO DE LOS ESTUDIOS INTERNACIONALES COMPARADOS DE RENDIMIENTO ESCOLAR: DATOS Y PRIMEROS INTENTOS DE INTERPRETACIÓN SOBRE LA BASE DEL ESTUDIO PISA

The informative value of the international compared studies of academic achievement: information and first attempts of interpretation on the base of PISA study



Eckhard Klieme

*Deutsches Institut für Internationale Pädagogische
Forschung, Francfort del Meno*

E-mail: klieme@dipf.de

Petra Stanat

Universidad Libre de Berlín

*Institut für Erziehungswissenschaften und Psychologie
Abt. Empirische Bildungsforschung*

E-mail: petra.stanat@fu-berlin.de

* Originalmente apareció bajo el título "Zur Aussagekraft internationaler Schulleistungsvergleiche. Befunde und Erklärungsansätze am Beispiel von PISA" en *Bildung und Erziehung*, 2002, vol. 55 núm. 1, págs. 25-44.

Resumen:

Partiendo de la reciente bibliografía acerca de TIMSS y de los nuevos datos de PISA, el presente trabajo se propone discutir algunos aspectos básicos de la investigación empírica comparada en materia de educación. Se sostiene que los objetivos primarios de las evaluaciones de rendimiento a gran escala son: (a) operacionalizar de las metas educativas y las competencias de los alumnos, (b) establecer modelos causales que sirvan para explicar los resultados del sistema escolar y (c) brindar una descripción de la eficacia de los sistemas e instituciones educativas. En los análisis comparados, se hace uso de un conjunto de indicadores que definen minuciosamente el perfil de los resultados a nivel nacional. Sin embargo, la interpretación y explicación de las diferencias en el plano internacional adolece todavía de problemas teóricos y metodológicos considerables. Ello es así porque la investigación empírica apenas está comenzando a entender la repercusión que tiene el contexto cultural en los resultados escolares.

Palabras clave: Alumnado - Evaluación institucional - Habilidades - Evaluación - Investigación educativa - Ciencias de la Educación - Educación Comparada - Institución escolar/educativa - Sistema educativo.

Abstract:

Based on recent literature about TIMSS and new results from PISA, the article intends to discuss basic issues of empirical comparative educational research. It is argued that large scale assessments primarily aim at (a) the operationalization of educational goals and student competencies, (b) causal models that explain the outcome of schooling, and (c) descriptions of the efficacy of educational systems and institutions. For comparative analyses, a set of indices is used which define a fine-grained profile of national results. Interpretation and explanation of inter-national differences suffer, however, from major theoretical and methodological problems. Especially, empirical research just begins to understand the impact of cultural context on school outcomes.

Key Words: Student - institutional Evaluation - Skills - Evaluation - educational Research - Sciences of the Education - Compared Education - school / educational Institution - educational System.

A partir de la década de 1990 y tras un intervalo de unos veinte años, han vuelto a realizarse en Alemania estudios de rendimiento escolar a gran escala. La razón de que este campo de estudios quedara fuera de los programas de investigación empírico-pedagógicos debe buscarse en el debate político y científico en torno a la evaluación de la escuela secundaria integrada o comprensiva (*Gesamtschule*) que se dio en los años 1970. No fue casual, por ello, que los nuevos impulsos provinieran del extranjero, por medio de estudios de la *International Association for the Evaluation of Educational Achievement* (IEA) como, por ejemplo, el Tercer Estudio Internacional de Matemáticas y Ciencias Naturales (TIMSS) y, más recientemente, el Programa Internacional de Evaluación de alumnos (PISA).

La comparación internacional ha generado –no sólo en Alemania, sino también en los EE. UU., donde los estudios de rendimiento escolar cuentan con una tradición mucho más sólida– una renovación del interés y una profundización en la toma de conciencia de los problemas (cf. SUTER 2000). El *National Center for Education Statistics* (NCES) lleva adelante, desde hace unos diez años, una estrategia de internacionalización de los estudios comparados de rendimiento, en la expectativa de obtener, por un lado, informaciones acerca del posicionamiento de los EE. UU. en la competencia internacional y de generar, por otro, propuestas de reforma educativa. Esta tendencia se continúa a nivel de las distintas regiones y distritos escolares de los EE. UU., que utilizan réplicas de los tests internacionales con el fin de compararse, en especial, con los países del Este asiático.

Si se observa con mayor detenimiento, puede apreciarse, sin embargo, que la lógica de la comparación internacional no ha alcanzado todavía el grado de desarrollo que sería

necesario y deseable para extraer de la comparación conclusiones político-educativas y pedagógicas. Las líneas de argumentación que típicamente se emplean en la discusión son de carácter más bien anecdótico. El alto nivel de rendimiento alcanzado por los alumnos del Este asiático se atribuye a la tradición confuciana; los buenos resultados del Japón se explican alternativamente por una enseñanza moderna y orientada a los problemas o bien por la influencia de las escuelas complementarias privadas; se destaca a Finlandia por el amplio reconocimiento y cualificación de sus docentes; el éxito relativo que los países anglosajones han obtenido en PISA se funda en los sistemas de verificación de calidad establecidos en dichos países desde hace años...

Por más reveladoras y sugestivas que puedan ser estas interpretaciones basadas en el análisis ideográfico de países, resultan poco productivas a la hora de un esclarecimiento de sistemático de las diferencias que pueda ser de utilidad para la elaboración de estrategias nacionales de acción. La investigación internacional empírica comparada en materia de educación se encuentra en la actualidad sólo en los comienzos del desarrollo de estrategias sistemáticas por medio de la cuales se puedan interpretar los estudios de rendimiento escolar con vistas a su utilización y aprovechamiento en la praxis educativa. A continuación se discutirán algunos elementos de una de esas estrategias y se los ilustrará por medio de los resultados provenientes del estudio PISA, en especial de su parte dedicada a la competencia matemática básica.

- En la primera sección se presentarán las posibilidades de aprovechamiento que se ofrecen, en una primera instancia, sin recurrir a marcos de comparación internacionales. También en este caso, en el marco puramente intranacional, los estudios de rendimiento escolar pueden ofrecer una idea de la estructura de la competencia de los alumnos, poner a prueba modelos de interpretación y dar lugar a comentarios y propuestas en el plano descriptivo.
- En la segunda sección se insertarán estas consideraciones en el marco internacional. Se establecerán índices para una serie de países. Estos índices determinan un criterio de calidad o un contexto funcional dentro de los sistemas educativos y conciernen –aparte del nivel de rendimiento medio– a aspectos relativos a la variación del rendimiento, a los resultados de los grupos extremos, etc. Son estos índices los que constituirán luego la base para la comparación descriptiva de los sistemas. Se formularán asimismo modelos interpretativos y de competencia de manera diferencial, es decir, de forma separada para cada país.
- En la tercera sección se tratarán diversos intentos de interpretación de las diferencias entre países. Tomando como referencia los informes PISA y la bibliografía empírica a la que ha dado lugar el análisis de los datos de TIMSS, se señalarán diversos intentos de interpretación.
- En la cuarta sección se discutirá, por último, la cuestión del papel que los condicionamientos culturales desempeñan en la interpretación de los niveles de rendimiento nacionales y su operativización empírica, basada en los resultados de los estudios de rendimiento escolar. Se considerará como tarea futura de la educación comparada llevar a cabo una reconstrucción teórica de la influencia que ejercen las culturas y los sistemas educativos, que debe construirse a partir de una comprensión profunda de los condicionamientos culturales y valerse, a la vez, de modelos estadísticos complejos.

1. La utilidad de la comparación de rendimientos escolares en el plano intranacional

En los años que siguieron a la publicación de los resultados de TIMSS (BAUMERT/LEHMANN et al. 1997) se ha llegado a comprender, también en Alemania, que los estudios de rendimiento escolar son, en esencia, algo más que unas “carreras de caballos”. Ello se debe a consideraciones que surgen de tres áreas, que pasaremos a exponer brevemente a continuación (cf. KLIEME/KÖLLER/STANAT 2001).

1.1. Objetivos de la educación y modelos de competencia

Como consecuencia del desarrollo e interpretación de procesos de evaluación, los estudios de rendimiento escolar obligan a que se presenten de forma abierta y se precisen los objetivos de la enseñanza y la formación escolares, de manera de promover su discusión racional. Por ejemplo, en el marco de la publicación del informe PISA de Alemania (BAUMERT/KLIEME/NEUBRAND et al. 2001), se da cuenta en detalle de por qué se escogieron para su estudio determinadas dimensiones de rendimiento –desde la competencia en lectura, pasando por la formación básica en matemáticas y ciencias naturales, hasta competencias interdisciplinarias tales como la autorregulación del aprendizaje y las habilidades de comunicación y cooperación–, de cuáles son los constructos pedagógicos, didácticos y psicológicos que se asocian con dichas dimensiones de rendimiento, qué relación guardan éstas con los planes de estudio y la praxis de la enseñanza en las escuelas (“validez curricular”), qué tipos de exigencias caracterizan la asignación de tareas y qué niveles de competencia pueden distinguirse. En el informe PISA de Alemania se hace, por último, un intento –basado en apreciaciones de expertos en materia de currículum y análisis didáctico– por definir estándares de rendimiento, es decir por determinar los niveles que se espera que alcancen las alumnas y alumnos en las dimensiones de rendimiento estudiadas. Es de este modo que los estudios de rendimiento escolar realizan un aporte significativo al discurso de los objetivos de la formación escolar, aun antes de publicar un informe acerca de un grupo, escuela o país en particular.

1.2. Modelos de interpretación

No menos importantes que los tests en sí resultan, en los estudios científicos de rendimiento escolar, los datos, recabados de forma paralela, de alumnos, clases y escuelas. Basándose en un diseño de investigación que sigue estrechamente las magnitudes de influencia conocidas en la investigación pedagógica empírica, se desarrollan cuestionarios amplios destinados a alumnos, docentes, padres y directores de escuela, cuyos datos se incorporan en modelos de interpretación complejos. Así, bajo la forma de modelos de análisis del camino (*path analysis*), el informe PISA alemán estudia, además, la influencia que tienen en los niveles de rendimiento alcanzados el entorno socioeconómico de los alumnos (posición económica y profesional de los padres, inquietudes culturales, el llamado “capital social”), las habilidades cognitivas básicas, las estrategias y la motivación del aprendizaje y el interés que los alumnos demuestran por determinadas materias, así como también, por último, otros aspectos de la enseñanza y la escuela (forma de la escuela y tipo de instrucción elegido, ambiente y recursos de la escuela, etc.). El punto central de PISA está en el registro diferenciado de los entornos familiares, mientras que el estudio precedente (TIMSS) investigaba de forma diferenciada las variables relacionadas con la enseñanza en clase (BAUMERT/LEHMANN et al. 2000). Desde el punto de vista de las ciencias de la educación, es

en estos modelos de interpretación donde reside la verdadera utilidad de los estudios de rendimiento escolar. A partir de ellos pueden derivarse inmediatamente conclusiones que sirvan como guía a posibles intervenciones. Estas conclusiones incluyen, entre otras, la cuestión de si el entrenamiento de estrategias de enseñanza, el fomento de la motivación, la incentivación de los alumnos en situación de desventaja social o la igualación de las diferencias sexuales pueden tener influencia en una mejora del nivel de rendimiento.

1.3. Descripción y comentarios

Finalmente, los resultados de los estudios de rendimiento escolar pueden entenderse también como un comentario, un informe dirigido en particular hacia una clase o escuela estudiada, pero también hacia el conjunto del sistema educativo. Es una práctica corriente en los estudios de rendimiento escolar informar a los establecimientos participantes acerca de los resultados obtenidos, así como también acerca de las condiciones sociales e institucionales en las cuales desarrollan su tarea. Por ejemplo, el consorcio PISA de Alemania, por un lado, ofrece informes a las escuelas que hayan tomado parte en el estudio y, por otro, remite un reporte detallado a los ministerios de Educación de los estados federados sobre el sistema educativo y los resultados de rendimiento. La calidad de tales comentarios depende, sobre todo, de los componentes de investigación mencionados antes en los puntos 1.1. y 1.2. Cuanto más diferenciadas sean las dimensiones de rendimiento reflejadas en los tests, tanto más diferenciado será el comentario. Si en las dimensiones de rendimiento se distinguen niveles de competencia de acuerdo a contenidos, el comentario puede también proporcionar datos concretos acerca de las tareas superadas y no superadas. Por último, el comentario puede contemplar asimismo otros factores de explicación que se hayan relevado en el estudio, como, por ejemplo, posición social y capacidades cognitivas básicas. Además de los resultados medidos en el test, se comunican también indicadores “depurados”, en los que estas características de entorno se registran en forma numérica (“*value added results*”).

La operativización de dimensiones y niveles de rendimiento, los modelos de explicación y propuestas de intervención, los comentarios acerca de los resultados y condicionamientos de los procesos educativos son datos que resultan relevantes para las ciencias de la educación, que son útiles para la política educativa y también significativos para la praxis pedagógica. Los estudios de rendimiento escolar cobran así valor informativo sin recurrir a la comparación internacional.

2. Más allá de los valores medios: indicadores para la comparación de sistemas educativos

La comparación internacional persigue, en última instancia, los mismos objetivos que los análisis intranacionales, pero ampliados en una perspectiva diferencial. Hemos de ilustrar nuestras apreciaciones mediante los datos y resultados del Programa Internacional de Evaluación de Alumnos (PISA: *Programme for International Student Assessment*), que se llevó a cabo por primera vez en 2000 bajo coordinación de la OCDE. En esa ocasión se sometieron a prueba entre 4.500 y 10.000 alumnos de 32 países, 28 de ellos miembros de la OCDE. La población objeto era el grupo de personas de 15 años. Las áreas en que se realizaron los tests fueron las de competencia en lectura, formación básica en matemática y ciencias naturales, así como también competencias interdisciplinarias, en especial aspectos del aprendizaje autorregulado (cf. OCDE 2001, BAUMERT/KLIEME/NEUBRAND et al. 2001). La OCDE repetirá este estudio con un ritmo de tres años. En el presente trabajo nos concentraremos en los

resultados obtenidos en el área de matemática que ofrece el informe PISA de Alemania (KLIEME/NEUBRAND/LÜDTKE 2001).

En las siguientes páginas se hará, en primer lugar, una descripción de los rendimientos comparados internacionalmente. Luego se expondrá la manera de estudiar diferencialmente la estructura de la competencia en matemática, es decir, la manera de establecer distintos perfiles para cada país. En último término se abordarán las diferencias entre países en vista de los factores condicionantes personales, sociales y escolares.

2.1. La distribución de rendimientos en la comparación internacional

La forma más corrientemente utilizada en los estudios de rendimiento escolar para la descripción de los resultados es la de determinar valores medios específicos para cada país. Esto hace posible establecer luego una clasificación jerarquizada o una estimación de intervalos. Por ejemplo, la evaluación de los tests de competencia matemática de PISA 2000 (KLIEME/NEUBRAND/LÜDTKE 2001) arrojó como resultado el hecho de que, en la escala seleccionada —en la que se ubicaba el valor medio general de los países miembros de la OCDE en 500 puntos, con una variación estándar de 100—, los resultados promedio de la mayor parte de los países miembros de la OCDE se situaban en una región en torno a medio punto de la variación estándar, por encima o por debajo del valor medio general (cf. Cuadro 1). Los puntos extremos los marcaban Luxemburgo (valor medio: 446) y los Países Bajos (564) —aunque este último estado fuera luego excluido de la evaluación debido a su escaso nivel de participación; su lugar pasó a ser ocupado por Japón (557)—. Alemania (490) y Suecia (510) se encuentran respectivamente por debajo y por encima de los valores de comparación dentro de la OCDE, en lo que respecta a intervalos estadísticamente relevantes. Entre los países participantes de PISA 2000, sólo México y Brasil (no representado en el Cuadro 1) alcanzaron valores claramente inferiores a la media.

Claro está que el valor medio nacional es un indicador excesivamente simplista para evaluar el funcionamiento del sistema educativo de un dominio dado. La forma de visualización escogida, tanto en el estudio PISA internacional (OCDE 2001) como en el informe nacional alemán (BAUMERT/KLIEME/NEUBRAND *et al.* 2001), permite dar cuenta, por medio de “percentiles”, de toda una serie de datos adicionales respecto de los rendimientos nacionales. De este modo, se puede representar, además de los valores medios, sus intervalos de confianza. Se advierte así con claridad que la diferencia de tres puntos en la escala de los tests matemáticos de PISA que puede verificarse entre Alemania y los Estados Unidos es de escasa relevancia estadística, ya que los valores de Alemania se encuentran dentro del intervalo de confianza de los Estados Unidos. La amplitud total de la barra es un indicador de la variación de los rendimientos escolares dentro de un país determinado. El extremo derecho representa el percentil 95, es decir, el punto en la escala del test que sólo ha logrado alcanzar o superar el 5% de los alumnos con mejores resultados en un país dado. De manera análoga, el extremo izquierdo indica el percentil 5, es decir, el valor que no ha llegado a alcanzar o superar el 5% de los alumnos con peores resultados. En el Cuadro 1, las mayores amplitudes se advierten en los casos de Grecia, Bélgica, la Federación Rusa y Alemania. Esto significa que Alemania pertenece al grupo de países con la mayor heterogeneidad de rendimiento escolar en el área matemática.

Un indicador interesante lo constituye no sólo el intervalo entre los percentiles 5 y 95, sino también su correspondiente ubicación. Por ejemplo, en el Cuadro 1 podemos advertir que los mejores resultados obtenidos por los jóvenes alemanes de 15 años en el área matemática se ubican a la misma altura que los mejores resultados obtenidos en los EE. UU.,

Dinamarca y Francia: en todos estos países el percentil 95 se ubica en la misma posición de la escala. Sin embargo, el percentil 5 de Alemania se ubica claramente por debajo del de los países antes mencionados. Esto quiere decir que Alemania se sitúa, en lo que respecta a los mejores resultados en el área matemática, más o menos en el valor medio de los países industrializados, pero también que enfrenta serios problemas en la zona de los resultados más bajos.



Imagen 1: Distribución de los resultados del test de matemática en los países participantes de PISA (Fuente: Instituto Max Planck de Desarrollo Humano). La franja media representa el valor promedio, mientras que las zonas grisadas a la derecha y la izquierda de la misma indican el intervalo de confianza (-/+ 2 error estándar respectivamente). Las zonas extremas están limitadas por los percentiles 5%, 10% y 25% (a la izquierda) y 75%, 90% y 95% (a la derecha).

Este tipo de análisis diferenciado de la distribución de rendimiento es de gran importancia para la investigación en las ciencias de la educación y la reflexión en materia de política educativa. El grado de la variación y los indicadores de los grupos extremos (los percentiles 5 y 95), sumados al valor medio, brindan una imagen detallada de la calidad y la homogeneidad de la competencia matemática con la que un sistema educativo “equipa” a distintos grupos de alumnos.

En tanto que la discusión pública suele centrarse en el tema de las diferencias entre los países, no debe olvidarse que las distribuciones en gran medida se superponen. Por ejemplo, en cualquier país de la Unión Europea (Alemania incluida), un alumno que haya alcanzado el promedio nacional quedaría dentro del 50% intermedio. En los países en desarrollo se verifica, en cambio, un marcado contraste. Ya México, país integrante de la OCDE, cae claramente —con un valor medio de 387 puntos— fuera del espectro de los restantes países de la OCDE. Y dado que la OCDE, en colaboración con el Banco Mundial, invita a participar en el estudio PISA a países de otras regiones —por ejemplo, Brasil en 2000 y China en la ronda siguiente—, PISA pasa a cobrar una influencia creciente en el control de la política educativa en otros países.

2.2. Modelos de competencia diferenciales: perfiles de rendimiento en distintos países

Pero más importante aún que la comparación de distribuciones resulta para la práctica pedagógica y la investigación de las ciencias de la educación —en particular, la didáctica especializada— el problema de identificar los puntos fuertes y débiles de las alumnas y alumnos de un país dado. Se intenta, por ello, establecer perfiles que contribuyan a identificar las áreas de rendimiento que requieren en especial de reformas.

Por ejemplo, del estudio PISA se extrae el hecho notable de que los países del Este asiático —Japón y Corea— se ubican, tanto en matemática como en el área de las ciencias naturales, media desviación estándar por encima de la media de la OCDE, mientras que en el terreno de la competencia en lectura, la diferencia se reduce a sólo un cuarto de la desviación estándar. Resulta evidente, entonces, que en el campo de la competencia en lectura ambos países no tienen tanto éxito como en la matemática y las ciencias naturales. (Probablemente la causa de ello esté en sus complejos sistemas de escritura.)

Pero los países pueden presentar también distintos perfiles incluso dentro de determinadas áreas de rendimiento o asignaturas escolares. Un posible motivo podrían ser las diferencias de acentuación de ciertos aspectos en la cobertura del currículum de la asignatura. A tal fin, en diversos estudios internacionales se pregunta a los docentes de los alumnos evaluados acerca de los contenidos de las tareas transmitidas en clase. Sobre esta base, Burstein (1989) ha mostrado que la acentuación de ciertos aspectos del currículum se refleja asimismo en el éxito del aprendizaje en las áreas respectivas. En los estudios de la IEA, *oportunity to learn* ha probado ser un factor de importancia a la hora de explicar los perfiles de rendimiento.

Los contenidos de las asignaturas no constituyen, sin embargo, el único aspecto que se refleja en los perfiles de rendimiento específicos de cada país. También la influencia de la “cultura de la enseñanza”, es decir, de las formas de enseñanza y aprendizaje difundidas en cada cultura, puede dar lugar a puntos fuertes y débiles. Para identificar tales perfiles, se diseñó, en el marco del TIMSS (KLIEME 2000, KLIEME/BAUMERT 2001), un método que fue aplicado recientemente también en el estudio PISA (KLIEME/NEUBRAND/LÜDTKE 2001). Lo

que se investiga es si las tareas que los alumnos resuelven de manera desproporcionadamente frecuente o infrecuente —en comparación con su nivel general de rendimiento— encierran además ciertas exigencias. Estas exigencias son las que representan, entonces, los puntos fuertes o débiles del país en cuestión.

En el test de matemática de PISA pudo establecerse así que los alumnos alemanes obtuvieron resultados relativamente altos, comparados internacionalmente, en los problemas que requieren para su solución de conocimientos y procedimientos matemáticos, mientras que su punto débil estuvo en la construcción de modelos para situaciones matemáticas más complejas. Se refleja en ello la llamada “orientación al cálculo” de la enseñanza de la matemática en Alemania, tantas veces criticada por los especialistas en didáctica. Es interesante observar que los alumnos estadounidenses —que alcanzan, en promedio, un nivel idéntico al de los alemanes— presentan un perfil exactamente inverso: su punto fuerte está en los ejercicios complejos de “generalización”, mientras que su punto relativamente débil se encuentra en el área de la “reproducción de conocimientos fácticos”. Resulta entonces evidente que, al menos en el perfil de rendimiento de los alumnos estadounidenses, puede verse el resultado de más de una década de profundas reformas en la enseñanza de la matemática, en tanto que en Alemania la reorientación hacia un aprendizaje tendiente a la comprensión y solución de problemas todavía está haciéndose esperar.

2.3. El acoplamiento de los rendimientos a las influencias personales, sociales y escolares

Para la planificación de intervenciones pedagógicas resulta de particular importancia, como ya se ha señalado en la Sección 1, identificar los parámetros de explicación de las diferencias en el rendimiento y, en la medida de lo posible, incorporarlos en modelos causales de pronóstico de los rendimientos en los tests. Ahora, bien, es probable que dichos modelos deban interpretarse de manera diferente en distintos países y culturas. Así, Scheerens y Bosker (1997, p. 255ss.) han mostrado, sobre la base de un estudio internacional de lectura, que, fuera de la variable “posición social de la familia”, no había en el estudio otro factor que pudiera tener una influencia general en todos los países participantes. Una conclusión básica que se obtiene de los numerosos estudios sobre la efectividad escolar es, por ejemplo, que en los países en desarrollo la disponibilidad de recursos materiales en la escuela puede generar grandes diferencias en los resultados, mientras que en los países industrializados la relevancia de este factor resulta mínima. Por el contrario, la calidad de la enseñanza tiene en los países industrializados una influencia particular que no se observa —por falta de variación en el estilo de enseñanza— en los países en desarrollo (SCHEERENS 2001). En el informe PISA internacional se incorporan modelos de interpretación para cada país (cf. OCDE 2001, tab. 8.5a), de los que se extrae que, también en este caso, pueden verificarse constelaciones muy variadas de país en país.

Desde la perspectiva de la educación comparada, esta estrecha relación entre el rendimiento escolar y los factores condicionantes constituye de por sí un indicador de particular interés. Por ejemplo, en varios países uno de los principales objetivos es el de disociar el éxito escolar del origen social, ya sea por consideraciones políticas de base democrática o en interés de la movilización de sus recursos humanos. Una correlación entre éxito escolar y origen social —obtenida por medios empíricos— puede ser de utilidad a la evaluación de tales objetivos. Unida con esta cuestión va también la medida en la cual el éxito de los alumnos se relaciona con el apoyo brindado por sus padres. De manera análoga, se puede plantear además el problema de las diferencias de rendimiento relacionadas

específicamente con el sexo. Por último, también es importante, desde el punto de vista de las ciencias de la educación, el papel que desempeñan los objetivos afectivos y cognitivos en la enseñanza. En este contexto, podría ser objeto de evaluación el grado de proximidad entre ambos tipos de objetivo; es decir, si existe un correlato entre, por ejemplo, el interés por la asignatura y el concepto personal que se tiene sobre el talento propio, por un lado, y el rendimiento de los alumnos, por otro. Desde la perspectiva de una diferenciación de estructuras individuales de la personalidad, sería deseable que pudiera establecerse un acoplamiento nítido, al menos en lo que respecta a la adolescencia. Como indicadores, podrían servir las correlaciones entre los rendimientos en los tests y el concepto personal que los alumnos tienen de su propio talento.

Sobre la base de los datos de PISA, se han calculado tales indicadores y se los ha comparado entre los países participantes.

- Los datos pertinentes a la posición socioeconómica de la familia de los alumnos se han recogido en PISA de manera especialmente precisa, válida e internacionalmente unificada. De tal modo se hace posible establecer de modo exacto la asociación existente entre los resultados y el origen social. Para ello puede utilizarse como índice (entre otros) la diferencia de los valores medios obtenidos por el 25% superior e inferior de la escala social (cf. BAUMERT/SCHÜMER 2001a, p. 385s.). En Alemania, esta diferencia en el terreno de la formación matemática básica es, con 98 puntos, superior a la de la mayor parte de los países restantes. Sólo en otros cuatro países —entre ellos, Suiza— puede encontrarse un acoplamiento más estrecho entre la formación matemática básica y el origen social. En el área de la competencia en lectura, Alemania es incluso el país que presenta el acoplamiento más estrecho entre rendimiento y posición social. Este dato —que se examina en detalle en Baumert y Schümer (2001a)— representa, en opinión de varios investigadores, el resultado de mayor trascendencia del estudio PISA en lo que a Alemania respecta, aunque demuestra también que, a más de cuarenta años del comienzo del proceso de expansión educativa, siguen imperando en el sistema educativo alemán las mismas disparidades sociales que antes.
- En 15 de los 32 países participantes se verifican asimismo significativas disparidades basadas en el sexo y, como era de esperar en vista de la bibliografía existente al respecto, siempre en favor de los jóvenes de sexo masculino. Como indicador del grado de esta desigualdad puede emplearse la diferencia de los valores medios alcanzados por muchachas y muchachos en la escala del estudio PISA de matemática. Esta diferencia es de 11 puntos en el promedio de la OCDE, de 15 puntos en Alemania, hasta alcanzar un valor máximo de 27 puntos en Austria y Corea (STANAT/KUNTER 2001). En el caso de Corea, la disparidad sexual alcanza valores similares también en el TIMSS (alumnos de 13 y 14 años de edad), si bien, en los países restantes, la comparación entre PISA y TIMSS en cuanto a la medida de las diferencias de rendimiento debidas específicamente al sexo exhibe una estabilidad sólo escasa.
- En todos los países participantes se verifica —de manera acorde con los datos que reiteradamente aporta la psicopedagogía— un acoplamiento estadísticamente significativo entre el rendimiento reflejado en los tests de matemática y el concepto individual que los alumnos tienen de su propio talento para la matemática. También en este caso puede establecerse un índice, que consiste en la correlación entre rendimiento y concepto individual. En Alemania, este índice es de un valor extremadamente bajo ($r = .089$). La relación entre concepto individual y rendimiento alcanza valores similarmente bajos (si se exceptúan ambos países latinoamericanos: México y Brasil) sólo en Suiza, Austria, Luxemburgo y los Países Bajos. Para la interpretación de este fenómeno debe tenerse en

cuenta la interacción entre el desarrollo individual y los contextos institucionales. Lo notable en este caso es que el sistema escolar de los países antes mencionados está segmentado según rendimiento. En la evaluación propia de sus conocimientos matemáticos los alumnos se orientan entonces a los marcos de referencia de su clase. De ahí que un “mal alumno” de un *Gymnasium* tenga un concepto individual de su rendimiento matemático más bajo que el de un “buen alumno” de una *Hauptschule*: este hecho puede servir para explicar la baja intensidad de la asociación.

- De manera análoga, hemos verificado la asociación entre el rendimiento matemático y el interés demostrado por la matemática como asignatura. También en este caso se obtiene un valor extremadamente bajo para los alumnos alemanes ($r = .089$), mientras que en Corea el acoplamiento entre rendimiento e interés es particularmente estrecho ($r = .349$).
- Otra variable de entorno social que hemos seleccionado es la del apoyo de los padres en el aprendizaje y las tareas escolares, que en los cuestionarios de PISA se refleja en una escala especial. En la mayoría de los países, la colaboración y apoyo de parte de los padres se da sólo en los casos de rendimiento insuficiente en matemática (¡!). Así ocurre en Alemania, donde el indicador —es decir, la correlación entre apoyo de los padres y rendimiento en matemática— es de $-.181$. Una excepción, en la que la correlación apunta en sentido contrario, la constituyen los países del Este asiático —Japón y Corea—. Resulta, entonces, evidente que en la mayoría de los países los padres se preocupan por el rendimiento escolar sólo cuando éste se ve de alguna manera comprometido, mientras que en los países del Este asiático los padres apoyan y fomentan a los alumnos precisamente en sus puntos fuertes. En este caso, el indicador de acoplamiento entre rendimiento y variable de entorno social puede interpretarse como signo de prácticas culturalmente específicas.

Los ejemplos arriba mencionados ilustran el hecho de que los estudios internacionales del tipo de PISA proporcionan una serie de indicadores que, más allá de la comparación de la distribución de rendimientos, ponen de manifiesto condicionamientos culturales y diferencias en el funcionamiento de los sistemas educativos. Ello no obstante, estos ejemplos dan lugar asimismo a toda una serie de problemas metodológicos. Así, nuestra ejemplificación relativa a las disparidades sexuales, en la que se observan diferencias entre los resultados de PISA y los de TIMSS, es indicio de que las relaciones bivariadas son, en cierto modo, inestables. Por otra parte, se dan otras circunstancias institucionales que dificultan la interpretación, como por ejemplo, los efectos que pueda tener un sistema educativo segmentado. La observación de que determinadas correlaciones —como en el caso de la relación entre el apoyo de los padres y el rendimiento— no sólo puede variar fuertemente en distintos países, sino también orientarse en sentido opuesto, pone en evidencia el hecho de que debe contarse con diferencias sustanciales en la dinámica de los sistemas educativos. Por todo ello, la interpretación de las diferencias establecidas entre los países —por no mencionar su explicación— se torna difícil y extremadamente condicionada. Para lograrlo, seguramente no basta con establecer simples correlaciones bivariadas.

3. Acerca de la interpretación de diferencias internacionales de rendimiento: la conexión de enfoques cuantitativos y cualitativos

En la sección precedente hemos observado algunas relaciones entre el rendimiento y distintos factores condicionantes dentro de los países y establecido que estas relaciones —debido, entre otras causas, a condicionamientos culturales— pueden tener diverso grado de intensidad o incluso orientarse en sentido opuesto. En las siguientes páginas se indagará una

posible explicación de tales diferencias entre países. La pregunta que ocupa el lugar central en el debate político –pero también en la discusión intradisciplinaria– es la de cuáles son los factores que permiten explicar la efectividad de los sistemas educativos, es decir, los niveles medios de rendimiento alcanzados.

Por ejemplo, la investigación de la influencia del apoyo de los padres se dirige, desde esta perspectiva, a la comprobación de si la media del apoyo de los padres de que informan los alumnos de un país dado se corresponde sistemáticamente con la media del nivel de rendimiento demostrado. Y, en efecto, en los 32 países participantes, esta correlación, con un valor de $-.06$, resulta poco significativa. Esto es señal de que existen relaciones que pueden demostrarse en el interior de los países y que, sin embargo, carecen de validez en un contexto internacional. De ello se desprende, como ya se ha señalado al final de la Sección 2.3, la directiva metodológica de escoger procedimientos analíticos que se ajusten a una estructura multinivel (es decir, a la articulación jerárquica alumnos-escuelas-formas escolares-sistemas educativos).

Esta aproximación –que sería deseable en sí– todavía no se ha puesto en práctica en la medida suficiente. Por ejemplo, Woessmann (2001) llevó a cabo un segundo análisis de TIMSS, en el que se identifican los exámenes finales centralizados a nivel nacional como característica de los sistemas educativos relacionada a un nivel de rendimiento alto. Sin embargo, sus análisis se limitan exclusivamente al nivel individual, razón por la cual se atribuyen a todos los alumnos por igual de un país dado ciertas características que conciernen al sistema educativo (por ejemplo, la forma y tipo de los exámenes finales). Somos, por ello, de la opinión de que sus resultados deben contrastarse en el marco de modelos multinivel más complejos.

El informe PISA internacional (OCDE 2001) hace uso de modelos de este tipo, en los que se incorporan magnitudes de influencia simultáneamente en diferentes niveles: a) variables individuales de los alumnos (entorno económico, social y cultural de la familia; sexo; condición de inmigrante; etc.), b) variables individuales de los alumnos agregadas a nivel de la escuela (nivel promedio de los entornos familiares) y c) variables relativas a la escuela (dimensiones de la escuela, proporción docente/alumno, recursos utilizados, aspectos de la atmósfera escolar y de la enseñanza), que se recaban en parte desde el punto de vista de la dirección de la escuela y en parte desde la perspectiva de los alumnos (y se agregan a nivel de la escuela). Es por medio de estas variables que se logra en el informe dar cuenta de aproximadamente un tercio de la varianza aritmética del rendimiento entre los distintos países (OCDE 2001, p. 205). La importancia de cada uno de estos factores, así como también del modo en que éstos interactúan entre sí, es algo que sólo podrá determinarse por medio de estudios más profundos, tales como los que planea la OCDE en el marco de las áreas temáticas adicionales.

Como ya lo han subrayado los autores en distintos pasajes del informe de la OCDE, resulta imposible interpretar de manera precisa los resultados de los análisis correspondientes a países que cuentan con un sistema escolar segmentado. Ello se aplica, por ejemplo, a los países de habla alemana y los estados del Benelux. En estos países se consigna a los alumnos, tras la escuela primaria, a distintas formas de escuela secundaria, dependiendo de su nivel general de rendimiento. Sin embargo, Baumert y Schümer (2000a) han demostrado que, además del nivel de rendimiento, es también el origen social de los alumnos lo que determina la selección del tipo de escuela. Es por ello que, en tales sistemas, las diferencias en el rendimiento de las escuelas del ciclo secundario se deben en parte al proceso de selección y en parte a las características de las escuelas. Si no se establece una clara distinción entre ambos efectos, es probable que se termine por sobrevalorar la influencia de las

características de las escuelas. En definitiva, es necesario proceder con suma cautela cuando se hace un análisis conjunto de sistemas educativos segmentados e integrados.

El informe PISA de Alemania (BAUMERT/KLIEME/NEUBRAND et al. 2001) se vale también de la comparación internacional para elaborar hipótesis acerca del desarrollo del rendimiento escolar en distintos estados federados alemanes. Estas suposiciones, sin embargo, no pueden corroborarse solamente dentro del marco que ofrecen los datos de PISA. Por ejemplo, para explicar la mejora en los rendimientos de matemática que los países anglosajones demostraron respecto del estudio TIMSS de los años 1994/95, se arguyen los intensos esfuerzos emprendidos en materia de control de estándares de calidad de la enseñanza y expectativas de rendimiento, pero también se hace mención de una acentuada tendencia hacia el planteamiento de problemas cercanos a la realidad, tales como los que se incluyen en el test de PISA (KLIEME/NEUBRAND/LÜDTKE 2001, p. 276). Otro esquema de argumentación que se utiliza en el informe PISA de Alemania consiste en poner en cuestión, mediante ejemplos en contrario, determinadas hipótesis. Por ejemplo, Schümer (2001, p. 416ss.) estudia la duración de los períodos lectivos en los países participantes (es decir, la cantidad de horas de clase por año escolar) y, tomando como base los casos de Islandia y Australia, llega a la conclusión de que las considerables diferencias observadas en la carga horaria no ejercen influencia alguna en las diferencias de rendimiento.

Llevado a sus últimas consecuencias, este tipo de argumentación conduce a una escrupulosa conexión entre los datos cuantitativos de los estudios de rendimiento escolar con los datos histórico-descriptivos de la educación comparada “clásica”. Sería de esperar que una mayor cantidad de conocimientos contextuales de los países participantes llevara a una mejor comprensión de las posiciones en la comparación internacional de rendimientos. Tras la publicación del estudio PISA, se han realizado algunos intentos al respecto (cf., por ejemplo, von KOPP/PRUCHA/GROLLMANN 2001). Pero, en tal caso, debe tomarse en cuenta un amplio espectro de factores, de distinta relevancia según el país en cuestión. Entre los que se mencionan recurrentemente, se cuentan: las tradiciones culturales –que se reflejan, por ejemplo, en la significación que se atribuye a la educación en la vida de los jóvenes y sus familias–, los condicionamientos económicos, las estructuras del sistema escolar (por ejemplo, estructura integrada, comprensiva o segmentada), la edad de ingreso a la escuela y la duración del ciclo escolar activo, la estructura, amplitud y profundidad del plan de estudios, las medidas de fijación de estándares y control de calidad, la estructura y calidad de la formación de docentes, las tradiciones didácticas y tipos de enseñanza culturalmente específicos y, por último, la conducta de los alumnos en su tiempo libre, sus actitudes hacia la enseñanza y sus estrategias cognitivas. Seguramente representaría un avance el hecho de que se dejase de producir informaciones sobre tal o cual factor particular en tal o cual país y se pasase de modo sistemático a informaciones que permitan una consideración contrastiva de varios sistemas. Ello significaría, además, todo un desafío para una educación comparada que tiende mayoritariamente a la mera descripción ideográfica.

Ahora bien, como se ha señalado antes, no es tarea sencilla “trasponer” informaciones ideográficas a modelos estadísticos. Tales modelos deberían operar de manera simultánea en varios niveles y ofrecer una interpretación multivariada, es decir, incluir diferentes magnitudes de influencia, cofundiciones y efectos de interacción. Y en el curso de un proceder de carácter puramente exploratorio, suelen producirse modelos estadísticos bastante poco “manejables” tanto en el aspecto técnico como en el teórico. Más prometedor resulta, al parecer, un procedimiento basado en modelos cuantitativos “sofisticados” y

estrictamente teóricos. Esto es lo que han venido llevando a cabo, en una serie de estudios, el sociólogo David Baker y sus colaboradores en Pennsylvania¹.

Tomando por base los resultados de TIMSS y sus estudios precedentes, han investigado los efectos de los currícula centralizados y descentralizados (STEVENSON/BAKER 1991), de la enseñanza coeducativa (BAKER/RIORDAN/SCHAUB 1995), de la enseñanza privada adicional y de apoyo escolar (“*Shadow education*”: BAKER/AKIBA/LETENDRE/WISEMAN 2001) y del tamaño de las clases (PONG/PALLAS 2001). Lo que hace que estos estudios resulten convincentes es su concentración en una de las problemáticas derivadas de la investigación de la calidad escolar. En estos casos, ya no se intenta explicar los datos estadísticos de la comparación internacional de rendimientos escolares de manera tentativa por medio de un “método de banda ancha” [*Breitband-Verfahren*], sino integrándolos intencionalmente en un diseño cuasiexperimental. Las variables estructurales relevantes (p.ej., en el estudio dedicado a la enseñanza complementaria, la cuestión de si ésta tiene carácter puramente auxiliar o si estimula potencialidades de excelencia) se contemplan por medio de indicadores inteligentemente definidos.

4. ¿Contextos culturales como parámetros de interpretación?

BOE, BARKANIC y BORUCH (2001) —a diferencia de Baker— han intentado llevar a la práctica el principio de análisis exploratorio hasta sus últimas consecuencias. En un extenso análisis de los resultados de TIMSS, han llegado a establecer la existencia de no menos de 170 variables de explicación potenciales. Por demás interesantes resultan las tres variables de mayor poder explicativo, por medio de las cuales —agregadas a nivel nacional— es posible dar cuenta de un 84% de la varianza en los rendimientos nacionales de los 41 países participantes en el área matemática. Se trata, en sentido estricto, no de variables relacionadas con la escuela o la enseñanza, sino con el grado en que alumnas y alumnos se han mostrado de acuerdo con las afirmaciones siguientes:

- a) “Soy bueno en Matemática” (correlación del rendimiento con el agregado a nivel nacional: $r = .60$).
- b) “Dice mi madre que es importante ser bueno en Educación Física” ($r = .76$)
- c) “Para ser bueno en Matemática se necesita suerte” ($r = .72$)

En los casos (b) y (c), es razonable esperar una correlación negativa (!) con el rendimiento en matemática, si se parte de la base de que un buen rendimiento en matemática supone dar prioridad a esta área sobre otros objetivos (p.ej., la educación física) y hacerse responsable del rendimiento personal propio. Y, en efecto, las afirmaciones (b) y (c) se corresponden, tanto a nivel nacional como a nivel individual, con resultados negativos en el área matemática. En el caso de la afirmación (a), se verifica lo contrario. Los resultados efectivos aumentan, por lo general, proporcionalmente al aumento en la valoración personal de alumnas y alumnos en un país dado respecto de su capacidad. En concordancia con BOE *et al.* (2001), puede decirse que este efecto remite al nivel de exigencia —culturalmente

¹ Véase los estudios posteriores de este autor sobre TIMSS, PISA y otras pruebas internacionales de rendimiento escolar: D. Baker y G. LeTendre, *National Differences, Global Similarities: World Culture and the Future of Schooling* (Stanford CA: Stanford University Press) y D. Baker y A. Wiseman (eds.) (2007). *Education for All: Global Promises, National Challenges*. (International Perspectives on Education and Society, Volume 8, Oxford UK: Elsevier Science Ltd). Nota de los compiladores de este volumen de *Profesorado*.

establecido— de los alumnos. En los países con niveles de exigencia altos, los resultados son mejores y las apreciaciones autocríticas de los alumnos más pronunciadas, sin importar el contexto causal de que se trate.

Las conclusiones obtenidas en BOE *et al.* (2001) son ejemplo de que la escala de los cuestionarios puede variar según el nivel de análisis. Observada en un alumno individual en comparación con sus compañeros, la afirmación (a) constituye, en principio, una información personal válida. Agregada a nivel nacional, la misma afirmación pasa a ser un indicador de valores nacionales. En última instancia, las tres afirmaciones, tomadas en conjunto y agregadas a nivel nacional, describen el valor que se asigna al esfuerzo personal y al rendimiento académico en el caso de la matemática. Y es posible incluso que se haya logrado identificar así ciertos condicionamientos culturales que definen el nivel de rendimiento de un país dado.

HOFER (2001) esboza una reflexión teórica respecto de estos hechos. Con el fin de explicar las diferencias en el rendimiento escolar internacional —a partir de TIMSS—, elabora una línea de argumentación que, partiendo de los valores culturales y los factores de motivación psicológica, conduce a los rendimientos escolares. El núcleo teórico de la argumentación, basada en la motivación psicológica, consiste en la hipótesis de que, en las sociedades postmodernas, el aprendizaje escolar enfrenta una creciente competencia de otras actividades a menudo más placenteras, de modo que resulta cada vez más difícil encontrar una motivación para el rendimiento. Basándose en una clasificación de los países participantes en TIMSS dividida según sociedades tradicionales, modernas y postmodernas (INGLEHART 1998), HOFER (2001, p. 5) pasa primero a comparar las sociedades tradicionales con las modernas: “El rendimiento escolar es mejor en los países en los que imperan valores modernos en comparación con los tradicionales. El resultado es casi trivial: significa sólo que los alumnos de los países con mayor PBI y presupuesto educativo obtienen mejores resultados”. Hofer procede luego a comparar las sociedades modernas con las postmodernas: “Con la dimensión de la postmodernidad (supervivencia contra búsqueda del bienestar) se produce [...] una correlación negativa en las sociedades industrializadas. En los países que presentan una proporción comparativamente alta de valores postmodernos, los rendimientos resultan peores”.

La tesis de Hofer parece también plausible a la vista de los datos del estudio PISA. La primera parte de su hipótesis se ve apoyada por el hecho de que los rendimientos obtenidos por México y Brasil son considerablemente inferiores a los del resto de los países participantes. En segundo lugar, la correlación —calculada supranacionalmente— entre el rendimiento en matemática y las exigencias de formación mencionadas por los alumnos mismos, su disposición al esfuerzo, la motivación instrumental extrínseca y el interés demostrado por la matemática es baja, con valores entre $-.46$ y $-.63$. A nivel intranacional, estas escalas de motivación se corresponden de modo absolutamente positivo con el rendimiento en matemática. Pero en forma agregada, no representan ya la motivación personal, sino un nivel, culturalmente establecido, de exigencia en el rendimiento. Su correlación con la motivación instrumental extrínseca sigue resultando significativa, con un valor de $-.40$, aun si se hace excepción de los casos de México y Brasil. Suponemos entonces que, en el plano nacional, cuando las informaciones personales de los alumnos implican un nivel de motivación extrínseca alto, ello puede ser indicio de que las alumnas y alumnos del país en cuestión no cuentan con niveles pronunciados de exigencia en el rendimiento y que, antes bien, poseen valores postmodernos. Con esta traslación de significados también se invierte su conexión respecto del rendimiento; es decir, cuanto más “postmodernos” —según nuestra investigación— sean los estándares y exigencias de rendimiento, tanto menor resultará —de acuerdo con la hipótesis de Hofer— el nivel de rendimiento medio. Ello puede

también ser indicio de la influencia fundamental que los condicionamientos sociales y culturales ejercen sobre la calidad de los sistemas educativos, si bien esta primera apreciación adolece del defecto, antes mencionado, de no dar cuenta de estructuras multinivel ni de los efectos ocasionados por las distintas formas de escuela.

A manera de conclusión puede afirmarse lo siguiente: el esclarecimiento sistemático de la calidad de los sistemas educativos resulta, en principio, posible dentro del marco de una comparación internacional que tome como criterio el rendimiento escolar. Las condiciones para ello son: (a) favorecer el acceso a un amplio espectro de informaciones contextuales provenientes de la educación comparada descriptiva; (b) incorporar al análisis condicionamientos culturales y estructuras de valores, desarrollando a tal fin indicadores cuantitativos; (c) aplicar modelos de análisis sofisticados, en los que (d) se brinde la debida atención a las estructuras escolares —en especial, en lo referente a los sistemas segmentados—. Son éstas expresiones de deseo que la investigación internacional comparada del rendimiento escolar todavía no ha llegado a hacer realidad.

Referencias bibliográficas

- Baker, D.P., Riordan, C., Schaub, M. (1995). The effects of Sex-Grouped Schooling on Achievement: The Role of National Context. *Comparative Education Review*, 39, n° 4, pp. 468-483.
- Baker, D.P., Akiba, M. Le Tendre, G.K., Wiseman, A.W. (2001). Worldwide Shadow Education: Outside-School Learning, Institutional Quality of Schooling, and Cross-National Mathematics Achievement. *Educational Evaluation and Policy Analysis*, 23, n° 1 (primavera), pp. 1-7.
- Baumert, J., Bos, W., Lehmann, R.H. (coords.) (2000). *TIMSS/III. Dritte Internationale Mathematik - und Naturwissenschaftsstudie - Mathematisch-naturwissenschaftliche Bildung am Ende der Schullaufbahn* (2 vol.). Opladen: Leske & Budrich.
- Baumert, J., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Stanat, P., Tillmann, K.-J., Weiß, M. (coords.) (2001). *PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich*. Opladen: Leske & Budrich.
- Baumert, J., Schümer, G. (2001a): Familiäre Lebensverhältnisse, Bildungsbeteiligung und Kompetenzerwerb. En: Baumert, J., Klieme, E., Neubrand, M. *et al.* (2001), pp. 323-407.
- Baumert, J., Schümer, G. (2001b). Lebens- und Lernbedingungen von Jugendlichen. Schulformen als selektionsbedingte Lernmilieus. En: Baumert, J., Klieme, E., Neubrand, M. *et al.* (2001), pp. 454-467.
- Boe, E.E., May, H.B.S, Barkanic, G.B.A., Boruch, R.F. (2001). *Predictors of National Differences in Mathematics and Science Achievement of Eighth Grade Students: Data from TIMSS for the Six-Nation Education Research Program. Research Report n° 2001-TIMSS2*. Filadelfia: Center for Research and Evaluation in Social Policy (Universidad de Pennsylvania).
- Burstein (1989). *Second International Study of Mathematics*, vol. 3. Londres: Pergamon.
- Hofer, M. (2001). Wertewandel und Schulmotivation. Pädagogische Implikationen postmoderner Werte. Conferencia pronunciada en el encuentro Konkurrierende Bildungsorte-Schule im Spannungsfeld von Familie, Jugendkulturen und Neuen Medien, Universidad Martín Lutero de Wittenberg-Halle, 7-9 de junio de 2001.
- Inglehart, R. (1998). *Modernisierung und Postmodernisierung*. Fráncfort del M.: Campus. [Modernización y posmodernización: el cambio cultural, económico y político en 43 sociedades. Madrid: Centros de Estudios Políticos y Constitucionales, 2006.]

- Klieme, E. (2000). Fachleistung im voruniversitären Mathematik - und Physikunterricht: Theoretische Grundlagen, Kompetenzstufen und Unterrichtsschwerpunkte. En: J. Baumert, W. Bos y R.H. Lehmann (coords.). *TIMSS/III. Dritte Internationale Mathematik- und Naturwissenschaftsstudie - Mathematisch-naturwissenschaftliche Bildung am Ende der Schullaufbahn* (vol. 2, pp. 57-128). Opladen: Leske & Budrich.
- Klieme, E., Baumert, J. (2001). Identifying National Cultures of Mathematics Education: Analysis of Cognitive Demands and Differential Item functioning in TIMSS. *European Journal of Psychology of Education*, XVI, n° 3, pp. 385-402.
- Klieme, E., Köller, O. Stanat. P. (2001). TIMSS und PISA: Von der Untersuchung fachlichen Lernens zur Analyse allgemeiner Kompetenzentwicklung. *Journal für Schulentwicklung*, 2, pp. 18-32 (??).
- Klieme, E., Neubrand, M., Lüdtke, O. (2001). Mathematische Grundbildung: Testkonzeption und Ergebnisse. En: Baumert, J., Klieme, E., Neubrand, M. et al. (2001), pp. 141-190.
- Le Tendre, G.K., Baker, D.P., Akiba, M., Goesling, B., Wiseman. A. (2001). Teachers' Work: Institutional Isomorphism and Cultural Variation in the U.S., Germany, and Japan. *Educational Researcher*, 30, n° 6 (pp. 3-15).
- Organización para la Cooperación y el Desarrollo Económicos (OCDE) (2001). *Knowledge and Skills for Life. First Results from the OECD-Programme for International Student Assessment (PISA) 2000*. París: OCDE.
- Pong, S.-L., Pallas, A. (2001). Class Size and Eighth-Grade Math Achievement in the United States and Abroad. *Educational Evaluation and Policy Analysis*, 23, n° 3 (otoño), pp. 251-273.
- Ramseier, E. (1999). Task Difficulty and Curricular Priorities in Science: Analysis of Typical Features of the Swiss Performance in TIMSS. *Educational Research and Evaluation*, 5, pp. 105-126.
- Scheerens, J., Bosker, R.J. (1997). *The Foundations of Educational Effectiveness*. Londres: Pergamon.
- Scheerens, J. (2001). Monitoring School Effectiveness in Developing Countries. En: Creemers, B.P.M., Reynolds, D. (coords.). *School Effectiveness and School Improvement*, 12, n° 4, pp. 359-384.
- Schümer, G. (2001). Lebens- und Lernbedingungen von Jugendlichen. Institutionelle Bedingungen schulischen Lernens im internationalen Vergleich. En Baumert, J., Klieme, E., Neubrand, M. et al. (2001), pp. 411-427.
- Stanat, P. & Kunter, M. (2001). Geschlechtsunterschiede in Basiskompetenz. En: Baumert, J., Klieme, E., Neubrand, M. et al. (2001), pp. 251-269.
- Stevenson, D.L., Baker, D.P. (1991). State Control of the Curriculum and Classroom Instruction. *Sociology of Education*, 64 (enero), pp. 1-10.
- Suter, L.E. (2000). Is Student Achievement Immutable? Evidence from International Studies on Schooling and Student Achievement. *Review of Educational Research*, 70, n° 4 (invierno), pp. 529-545.
- von Kopp, B., Prucha, J., Grollmann, P. (2001). *Trends in Bildung-international*. 1ª edición. Fráncfort del M.: Deutsches Institut für Internationale Pädagogische Forschung.
- Weiß, M., Steinert, B. (2001). Lebens- und Lernbedingungen von Jugendlichen. Institutionelle Vorgaben und ihre aktive Ausgestaltung - Die Perspektive der deutschen Schulleitungen. En: Baumert, J., Klieme, E., Neubrand, M. et al. (2001), pp. 427-454.
- Woessmann, L. (2001). Why students in some countries do better: International evidence on the importance of education policy. *Education Matters*, pp. 67-74.