

5



DECONSTRUYENDO LOS PUENTES DE PISA: DEL ANÁLISIS DE RESULTADOS A LA PRESCRIPCIÓN POLÍTICA

***Deconstructing PISA's Bridges: From Results
Analysis to Political Prescriptions***

Francesc Pedró*

RESUMEN

Los estudios PISA tienen un impacto destacable en la configuración de políticas educativas, aún a pesar de algunas limitaciones tanto en términos de opciones metodológicas como de su modelo explicativo. Cada vez más los estudios PISA demuestran cierta ambición prescriptiva a la que se da salida ya sea influyendo sobre las definiciones curriculares o bien por medio de la abierta definición de las políticas que, supuestamente a partir de las evidencias halladas en PISA, pueden contribuir a mejorar los aprendizajes de los estudiantes de forma significativa. Ambos mecanismos son analizados, al mismo tiempo que se sugiere una agenda de investigación precisamente destinada a conseguir que los estudios PISA puedan apoyar mejor los procesos de configuración de políticas educativas basados en evidencias.

* UNESCO

PALABRAS CLAVE: PISA, Políticas Públicas, Políticas Educativas Nacionales.

ABSTRACT

PISA studies are having an impressive impact on educational policy making, despite some constraints both in terms of methodological choices and of its explanatory model. Increasingly PISA studies show a prescriptive ambition, which is fulfilled either by influencing curriculum definitions or by openly defining which policies can contribute to improve students' attainment significantly, supposedly drawing on PISA evidence. Both mechanisms are analyzed while at the same time a pending agenda for further research intended to make PISA studies more supportive of evidence-based policy making is suggested.

KEY WORDS: PISA, Public Policies, National policies on Education.

INTRODUCCIÓN

A estas alturas no parece necesario insistir en la importancia de los estudios PISA (*Programme for International Student Assessment*) de la OCDE¹ (Organización para la Cooperación y el Desarrollo Económico) y el impacto que tienen en las discusiones acerca de la política educativa, tanto a escala nacional como internacional, muy por encima de otros estudios comparativos semejantes como, por ejemplo, TIMSS (*Trends in Mathematics and Science Study*) y PIRLS (*Progress in International Reading Literacy Study*) de la IEA² (*International Association for the Evaluation of Educational Achievement*), o TERCE (Tercer Estudio Regional Comparativo y Explicativo del Aprendizaje) de la UNESCO³. También parece fuera de toda duda que PISA cuenta con una metodología, unos instrumentos de medida y unos controles de calidad apropiados a sus finalidades. La comunidad científica aprecia igualmente la accesibilidad a los datos y la transparencia de los análisis aplicados, puesto que se hacen públicos sin limitaciones para permitir así

¹ Para tener mayor información acerca de los objetivos del Programa, su metodología y resultados, así como para acceder a los datos e incluso realizar análisis en línea, véase www.oecd.org/pisa.

² Para tener mayor información sobre ambos proyectos, véase www.timss.org.

³ Más detalles sobre este Programa y sus antecesores en América Latina en www.llece.org.

una verificación externa, al tiempo que la reutilización de los datos para análisis ulteriores.

Pero, los estudios PISA fueron diseñados para ir mucho más allá de la descripción comparativa de los logros de los alumnos de 15 años de edad y del análisis de sus determinantes y condicionantes, tanto a escala sistémica, de centro o de alumno. Su ambición de partida es la de arrojar luz sobre las políticas públicas de educación, resaltando especialmente aquellas que conducen a la mejora de los resultados de los alumnos. Para cumplir con esta ambición, los estudios PISA establecen una serie de puentes desde el análisis de los resultados hasta la prescripción de políticas públicas en educación cuyo entramado sólo puede ser apreciado si es deconstruido por piezas. Pero, ¿se trata verdaderamente de puentes o, más bien, de saltos en el vacío? Dicho de otro modo, ¿las lecciones políticas que se predicán invocando los datos de los estudios PISA están suficientemente basadas en evidencias empíricas?

Responder a estos interrogantes es, precisamente, el objetivo de esta contribución. Para hacerlo en un contexto adecuado, la contribución se estructura en tres apartados. El primero intenta dar buena cuenta del impacto de los estudios PISA en los debates sobre las políticas nacionales de educación, demostrando que es extremadamente notable en algunos de ellos. El segundo apartado presenta algunas de las limitaciones metodológicas de estos estudios, haciendo especial hincapié precisamente en la ambición prescriptiva de los análisis y en su fundamentación empírica. El tercero establece lo que podría considerarse como una posible agenda de investigación para el futuro, sino para los estudios PISA cuando menos para los estudios comparativos internacionales basados en la evaluación de los logros de los alumnos si efectivamente se desea que puedan arrojar luz sobre las políticas públicas que pueden contribuir decisivamente a la mejora de los aprendizajes de los alumnos.

1. UN INNEGABLE IMPACTO POLÍTICO

Es indudable que los estudios PISA no tienen como objetivo la mera ordenación de los países participantes en un *ranking*, sino informar las políticas educativas poniendo de manifiesto cuáles son los determinantes y los

condicionantes de los aprendizajes de los alumnos, ya sea por la vía de las correlaciones con variables sistémicas, de centro o incluso del alumno a escala individual o agregada. Lógicamente, si se llega a un conocimiento de base empírica acerca de la forma en que variables independientes, que son eventualmente modificables por medio de intervenciones, inciden en el aprendizaje de los alumnos, entonces se entra de pleno en el terreno del diseño de políticas públicas basadas en la evidencia (OECD, 2007a; Slavin, 2008). Pero el poder de los *rankings* generados a partir de los datos de PISA es tal que su mera publicación se convierte en el detonante de enconadas discusiones políticas a escala nacional. En definitiva, aunque sea como resultado a veces de una lectura muy superficial, los estudios de PISA tienen ya un impacto sobre las políticas educativas. Este impacto se produce por dos vías: mediática y discursiva.

Por una parte, la presencia mediática de los resultados es impresionante en todos los países participantes. Pero, como se ha mostrado en distintos casos esta presencia es superficial y amplifica todo cuanto arroja una imagen negativa del sistema escolar del país. En efecto, la presencia mediática contribuye enormemente a generar la imagen pública del sistema escolar de un país en la opinión pública y, generalmente, su efecto es devastador. En una comparación acerca de la recepción por parte de la prensa de los resultados de PISA 2006, en Alemania, Francia y Gran Bretaña, Dixon et al. (Dixon, Muller, Arndt, & Hood, 2008) han demostrado que la negatividad es la característica dominante en la prensa de todos estos países, aunque con algunas diferencias. Así, por ejemplo, los artículos británicos y franceses fueron siempre mucho más negativos que los que aparecieron en la prensa alemana, pero en términos relativos el porcentaje de artículos críticos con los resultados nacionales fue siempre significativamente superior al de los artículos que mostraron neutralidad o fueron abiertamente negativos con el propio sistema escolar. El mismo estudio demostró igualmente que la receptividad a este tipo de estudios varía también de acuerdo con el país: Alemania parece mu-

⁴ La expresión "shock de PISA" se refiere a la conmoción pública creada por la publicación de los resultados de la primera oleada de PISA (2000) que, en cierta medida, ya habían sido adelantados años antes por los resultados de TIMSS. En el ranking resultante de PISA 2000 Alemania ocupaba la posición número 23 de un total de 30 países participantes. Para un análisis más detallado del shock de PISA véase, por ejemplo, el documentado trabajo de Gruber (Gruber, 2006).

cho más proclive a exorcizar los resultados desde la época del shock de PISA⁴ que Francia o Gran Bretaña y, de hecho, tras la difusión de los estudios de la oleada de 2006, se publicaron 15 veces más artículos sobre PISA en la prensa alemana que en la de los otros dos países en el mismo lapso de tiempo (de noviembre a diciembre de 2007), mostrando efectivamente que el shock de PISA seguía dominando el interés de los medios alemanes en materia educativa. Parecida evaluación se ha llevado a cabo con el reflejo en la prensa canadiense y los resultados son consistentes con esta idea de que lo más atractivo para los medios es lo más crítico para la situación relativa del país (Stack, 2006).

Por otra parte, la presencia discursiva en política educativa es igualmente destacable y, en cierta forma, es inducida también por la recepción mediática. La publicación de los resultados de una oleada PISA se traduce en un posicionamiento inmediato por parte de los actores relevantes en política educativa, ya se trate del gobierno, la oposición, los sindicatos o las organizaciones empresariales, así como de las familias. Pero, además, se espera que cualquier programa o estrategia política que tenga que ver con la educación obligatoria incluya una mención específica a los resultados de PISA (Grek, 2009). De hecho, en el discurso político a escala nacional parece ser más frecuente la referencia a los estudios PISA que a los resultados de las evaluaciones nacionales cuando lo razonable sería lo contrario, tal y como se ha demostrado en países muy distintos en cuanto a su posición relativa en el ranking de PISA, desde Alemania hasta Canadá, pasando por Francia, Inglaterra o Suiza (Hopmann, Brinek, & Retzl, 2007), aunque con la notable excepción de los Estados Unidos (Schneider, 2009) donde la recepción de los estudios PISA o TIMSS tropieza con una clase política mucho más interesada por la discusión de los diferenciales de resultados entre los distintos estados en las pruebas nacionales como las de NAEP (*National Assessment of Educational Progress*)⁵. Claramente, esta cuestión remite a la más amplia de las lecturas críticas de los estudios PISA, que constituye el eje central de esta contribución.

⁵ Se trata de una prueba de evaluación que se administra a muestras representativas de alumnos de 9, 13 o 17 años de edad en un amplio espectro de áreas curriculares que van desde la escritura, las matemáticas y las ciencias (como en los estudios PISA) hasta el arte, la educación cívica, la economía, la geografía o la historia de los Estados Unidos. Para más detalles véase nces.ed.gov/nationsreportcard.

2. LOS ESTUDIOS PISA TAMBIÉN TIENEN LIMITACIONES

Existe un consenso generalizado acerca de la solidez metodológica de los estudios PISA, desde la concepción del marco de evaluación y la definición de las competencias hasta los procedimientos de control de calidad en todo cuanto se refiere a los procedimientos de examen. Esta solidez viene reforzada, por otro lado, por la transparencia tanto de los cálculos como de los análisis, con la idea de que puedan ser replicados por observadores externos cuantas veces haga falta.

Con todo, este consenso generalizado junto a la curva de aprendizaje exigida por la complejidad tanto del diseño metodológico como de las estrategias y técnicas de análisis, ya sean éstas psicométricas o econométricas, hacen más difícil que los usuarios directos de los estudios PISA lleguen a tener cierta comprensión de sus limitaciones. El resultado es que hay un cierto número de críticas que no son frecuentemente evocadas. Algunas de ellas son estrictamente limitaciones u opciones técnicas y sugieren que existirían alternativas a las estrategias aplicadas. Otras, tal vez las más importantes, tienen que ver con cuestiones mucho menos técnicas y más profundas, aunque están igualmente relacionadas con las opciones metodológicas y, en particular, con el modelo explicativo. La ausencia de los procesos de enseñanza y aprendizaje en la ecuación explicativa de los logros de los alumnos resulta, bajo este punto de vista, difícil de justificar.

2.1. Limitaciones derivadas del diseño metodológico

Las discusiones acerca de las limitaciones técnicas de los estudios PISA acostumbran a ser prolijas en aspectos relacionados fundamentalmente con la construcción de las pruebas y de los distintos ítems. Pero, si se acepta el principio de que los estudios PISA tienen como objetivo la medida de las competencias de una muestra representativa de los alumnos de 15 años de edad de cada país participante en lenguaje, matemáticas y ciencias, ninguna de las críticas parece sugerir que pueda existir en conjunto una estrategia técnica mejor.

Así, por ejemplo, en términos de validez técnica del estudio se han mencionado con cierta frecuencia algunos problemas que tienen, en realidad,

difícil solución como, por ejemplo, el del control de las tasas de respuesta (Prais, 2003) o la elección de la población basada en un criterio de edad y no de curso (Chromy, 2002), de modo que el universo del que se extrae la muestra es la población de alumnos de 15 años de edad y no, por ejemplo, los alumnos que están cursando el cuarto curso de la enseñanza secundaria obligatoria. Pero, quizás los dos aspectos que merecen mayor atención, por sus eventuales implicaciones para la interpretación política, son la motivación de los alumnos y el riesgo de que se les pueda entrenar para superar este tipo de pruebas.

Como ha señalado Martínez Arias (Martínez Arias, 2006), aún queda por resolver un problema importante de las evaluaciones a gran escala, y es el de la motivación de los alumnos cuando participan en evaluaciones que no tienen consecuencias para ellos como las tendría, por ejemplo, un examen de reválida⁶. O, lo que también es posible, sólo los alumnos motivados se prestan a pasar la prueba⁷. La investigación sugiere que cuando no hay consecuencias para ellos los alumnos tienden a no implicarse suficientemente en la resolución de las tareas (O'Neil, Abedi, Miyoshi, & Mastergeorge, 2005; Wise & Demars, 2003) y, como consecuencia, las puntuaciones de las pruebas podrían no ser indicadores válidos de sus competencias.

El problema puede ser más grave en las evaluaciones internacionales, dado que podrían producirse interacciones con las diferentes culturas de los países implicados. Aunque es cierto que se presta mucha atención a la traducción de los ítems a los idiomas correspondientes y a la equivalencia cultural, algunos observadores han señalado con perspicacia que existe una orientación occidental y urbana en quienes diseñan las pruebas (Dall, 2010). Por ejemplo, en la prueba PISA 2000 se solicitaba la interpretación de un *graffiti*, algo frecuente en las grandes ciudades europeas o norteamericanas, pero ciertamente despreciable para una joven de Singapur, donde el *graffiti* puede ser penado con ocho golpes de bastón o hasta tres años de cárcel (Kolesnikov-Jessop, 2010). Además, particularmente a los 15 años de edad,

⁶ En inglés se distingue entre *high stakes test* y *low stakes test* para diferenciar así los grados de interés que el sujeto puede tener por conseguir un buen resultado y su consiguiente motivación. Obviamente, en este sentido PISA es un *low stakes test*.

⁷ Este parece haber sido el caso de los Estados Unidos en la oleada de PISA 2006, donde hubo problemas de diversa índole, pero en particular para conseguir el número crítico de participantes (Schneider, 2009).

el grado de desafección de los alumnos con respecto a la escuela puede ser muy diferente en distintos países, como ya ha demostrado la investigación comparativa internacional (Danielsen, Samdal, Hetland, & Wold, 2009). Como se afirma en un informe de validación (Ruddock, T., Purple, & Ager, 2006), «en PISA, la buena disposición de los alumnos a leer el volumen de texto requerido, y su capacidad para hacerlo, son probablemente los dos factores cruciales» (p. 95). Puede esperarse, por consiguiente, que a mayor desafección, menor interés por responder adecuadamente a una prueba sin consecuencias prácticas para el alumno.

El segundo aspecto que podría ser eventualmente preocupante desde una perspectiva política es la posibilidad de entrenamiento previo a la administración de las pruebas. Lógicamente, una de las obvias limitaciones de cualquier tipo de prueba externa codificada es que siempre existe la posibilidad de ejercitar a los individuos para enfrentarse a ella. Esto, por lo menos en teoría, sugeriría que los resultados obtenidos pueden estar contaminados por una preparación previa de los sujetos. Algunas voces han sugerido, aunque nunca publicado en revistas académicas, que los resultados en PISA de algunos países, particularmente en el continente asiático, pueden deberse precisamente a un entrenamiento previo. Pero, la verdad sea dicha, nadie ha conseguido demostrarlo y es difícil que se logre por la sencilla razón de que las pruebas PISA están diseñadas con el objetivo de medir competencias y no conocimientos; dicho de otra manera, la única forma en la que se podría entrenar a los alumnos para mejorar sus resultados en pruebas de competencias sería, precisamente, formarles en esas competencias. Todo esto no deja de ser paradójico, pero ciertamente posible como la investigación empírica ha demostrado (Brunner, Artelt, Krauss, & Baumert, 2007). Pero, ¿sería deseable? Bien está lo que bien acaba y, por consiguiente, bienvenida sea cualquier estrategia que efectivamente se traduce en logros de los alumnos, pero cabe preguntarse acerca de los beneficios de unos procesos de enseñanza y aprendizaje orientados exclusivamente a la superación de pruebas de evaluación.

Otra cosa bien distinta es si la presión evaluadora a la que están sometidos los alumnos, y la forma en que responden a ella, en cada país es equivalente o no. Dicho de otra forma, los alumnos de países donde la presión evaluadora externa es elevada (por ejemplo, en forma de reválidas o de ejercicios corregidos por equipos docentes distintos a los del propio centro escolar), estarían mucho más habituados a enfrentarse a pruebas y serían ca-

paces de solventar mucho mejor el estrés vinculado al examen. No hay estudios todavía que examinen empíricamente esta cuestión para una muestra amplia de países⁸, pero la mera observación permite concluir que en países como Finlandia, Corea o China (Shanghai, Hong Kong) los alumnos están mucho más entrenados para enfrentarse a pruebas externas que, por ejemplo, en España. Muy probablemente la presión evaluadora, y la forma en que los alumnos responden a ella, estén lejos de ser el conjunto de factores más importante en la explicación de los resultados de los alumnos españoles, pero eso no significa que no contribuyan a ellos de una u otra forma.

2.2. Limitaciones derivadas del modelo explicativo

Cualquier docente o directivo escolar que se acerque al modelo explicativo de los estudios PISA se encontrará con la sorpresa de que los procesos de enseñanza y de aprendizaje, por lo menos por el momento, no tienen ningún lugar asignado. En suma, a todos los efectos, el aula es considerada una «caja negra» y, por consiguiente, el mensaje que se transmite tanto a la opinión pública como a los decisores políticos es que la pedagogía y, más en concreto, la didáctica son ni más ni menos que irrelevantes.

La justificación, si cabe, de esta ausencia tiene probablemente dos argumentos: uno derivado de la naturaleza del modelo explicativo y otro de las dificultades metodológicas inherentes al estudio empírico de los fenómenos dentro del aula con vocación representativa.

El primer argumento se basa en el hecho de que la función de productividad en educación no contiene ninguna referencia a los procesos de enseñanza y aprendizaje, probablemente porque (véase el segundo argumento) la dificultad heurística es tan grande que se prefiere prescindir de esta variable. En efecto, en la literatura econométrica la denominada función de producción en educación es relativamente popular. En esencia, en su definición más clásica (Hanusek, 1979; Levin, 1974), se representa de la forma siguiente:

$$A_{ij} = f(S_{ij}, B_{ij}, P_{ij}, I_{ij})$$

⁸ Aunque sí para parejas de países como, por ejemplo, Dinamarca y Finlandia (Andersen, 2010).

Donde A_{ij} representa el resultado del proceso educativo del alumno i en la escuela j , que depende de un conjunto de factores representados por los inputs escolares (S_{ij}), las características socioeconómicas de los estudiantes (B_{ij}), las influencias de los compañeros o efecto del grupo de pares (P_{ij}) y la capacidad innata del alumno (I_{ij}). El resultado u *output* en los estudios PISA está representado por el resultado obtenido por los alumnos en una prueba estandarizada. Este inevitable reduccionismo (de los resultados de la educación a resultados de pruebas objetivas), y cuyo uso está ampliamente respaldado por la literatura especializada ([Hanusek, 2003](#)), está igualmente aceptado como referencia por los actores principales de la educación. Nada dice la ecuación de cuanto acontece en el aula; es más, no necesita referirse a ella para explicar una parte significativa de la varianza de los resultados. Increíble, pero tan cierto como triste pedagógicamente hablando.

El segundo argumento sostiene, con razón, la dificultad intrínseca de introducirse en el aula para analizar lo que sucede en su interior. En particular, se trata de: a) documentar las prácticas de enseñanza y aprendizaje; b) hacerlo de forma estadísticamente representativa de la varianza en métodos que se puede dar en un país, a lo largo de todo el itinerario escolar hasta que los alumnos cumplen 15 años de edad; y c) establecer una conexión entre estas variables y los logros de los alumnos, lo cual es probablemente lo más difícil, no ya de conseguir, sino incluso de imaginar de forma operativa.

Por consiguiente, las dificultades metodológicas existen, pero se podría hacer bastante más. Habida cuenta de estas dificultades quizás haya que marcarse un horizonte poco pretencioso y limitarse, para empezar, a documentar las prácticas predominantes y a explorar métodos de codificación de las observaciones suficientemente robustos como para entrar en la ecuación explicativa. De sobras es sabido, para empezar, que en el ámbito de los estudios comparativos de educación existe un notable acervo de observación directa y análisis de las prácticas de enseñanza y aprendizaje en el aula ([Baker & LeTendre, 2005](#); [Brint, 2006](#); [Osborn et al., 2003](#)) que, por cierto, los estudios PISA nunca han tomado en consideración. Para PISA, sin embargo, lo que acontece en el aula sigue siendo un misterio, excepto por las percepciones de alumnos, profesores y directivos expresadas por medio de encuestas sobre el clima escolar, por ejemplo.

En efecto, el aula es, a todos los efectos, una «caja negra» para PISA. En el fondo, se puede argumentar que se trata de una cuestión esencialmente metodológica habida cuenta de las dificultades que entrañaría un análisis de las prácticas pedagógicas en el aula que realmente pudiera dar cuenta de la variabilidad existente dentro de un mismo país, por no decir dentro de un mismo centro escolar, representándolas apropiadamente. Pero, en segundo lugar, más difícil sería aún vincular prácticas con resultados escolares en el contexto de PISA. De hecho, expresado así sería inviable puesto que los estudios PISA evalúan los logros de aprendizaje conseguidos por los alumnos de 15 años de edad de forma acumulativa a lo largo de su vida, es decir, no en relación a un curso académico o a una materia concreta sino limitándose a medir lo que los alumnos, en un determinado momento de su vida, pueden acreditar en relación a las competencias que se quieren evaluar sea cual sea la forma en que las hayan adquirido.

Sin embargo, parece claro que hay razones tanto de tipo pedagógico como político para que la agenda de la investigación en PISA, o en los estudios satélite como TALIS⁹ (*Teaching and Learning International Survey*) (OECD, 2009), incluya una mayor transparencia de cuanto acontece en el aula. De hecho, existen interesantes precedentes metodológicos en TIMSS, utilizando las grabaciones de video de sesiones de clase como punto de partida para un análisis de las prácticas de enseñanza y aprendizaje (Janik & Seidel, 2009).

Desde un punto de vista pedagógico, la ausencia del aula en la ecuación explicativa de los resultados escolares contradice no sólo las expectativas de los profesionales de la educación sino, abiertamente, el sentido común. ¿Se pueden llegar a explicar los resultados que obtienen los alumnos de un país sin que se considere para nada el interior del aula? ¿Son suficientes los elementos contextuales, políticos y estructurales de los sistemas educativos para explicar los resultados de los alumnos? ¿Nada tienen que ver con los resultados la forma en que los docentes enseñan y los alumnos aprenden?

Desde un punto de vista político, la principal implicación de la desconsideración del aula es, ni más ni menos, que un desprecio implícito por el trabajo de los docentes. Si el mensaje que llega a la opinión pública es el de

⁹ Para mayor información sobre este Programa, véase www.oecd.org/talis.

que los resultados del sistema escolar son, por ejemplo, peores que los de la media de los países de la OCDE, sin que se pueda predicar absolutamente nada de la forma en que los docentes realizan su trabajo en el aula, el mensaje implícito es doble. Al hacerlo, se está afirmando, por una parte, que vistos los resultados de los alumnos los docentes no hacen bien su trabajo y, por otra parte, que las estrategias pedagógicas son prácticamente irrelevantes. Esto deja al profesorado no sólo sin argumentos acerca de su profesionalidad sino, lo que todavía es peor, sin una idea clara acerca de cómo podrían contribuir, tal vez haciendo las cosas de otra forma, a mejorar el desempeño de los alumnos. Y, lo que es más, olvida algo obvio: que cualesquiera que sean las reformas políticas propuestas nada cambiará en el aula si los docentes no lo hacen posible (Kenney, 2005; Smith, 2008). Algo, paradójicamente, que la propia OCDE se afana en recordar últimamente (OECD, 2010a, 2011).

3. CRUZANDO LOS PUENTES HACIA LA PRESCRIPCIÓN

Muy probablemente el sueño de todo científico social, por lo menos desde Popper, es el de contribuir a la solución de problemas por la vía de la prescripción basada en la evidencia científica. Desde Holmes, los estudios comparativos de educación no han escapado a esta noble ambición. Y los estudios PISA no son una excepción puesto que intentan construir puentes desde los datos comparativos internacionales hacia la prescripción nacional. PISA lo hace de dos formas distintas: por una parte, en lo que probablemente sea un puente invisible para el que, sin saberlo, lo está cruzando, genera prescripciones curriculares y, por otra parte, en lo que seguramente es el puente más largo y arduo de cruzar, prescribe políticas públicas de educación que deberían mejorar los resultados de los alumnos.

Pero, ¿cuentan estos dos puentes con suficientes fundamentos? Para comprobarlo habrá que deconstruirlos.

3.1. El puente invisible: la prescripción curricular

La confusión acerca de lo que miden los estudios PISA sigue siendo, a pesar de todos los esfuerzos, muy importante. En una aproximación muy

simplista se trata de evaluar lo que saben los alumnos de 15 años de edad, desde una perspectiva comparativa internacional, tomando una muestra representativa de cada país participante. Pero esta aproximación simplista no reconoce algo que es fundamental: lo que los estudios PISA intentan medir no son los conocimientos de los alumnos sino sus logros en relación a una escala de competencias definida previamente por el propio Programa. Es importante insistir en que esta escala de competencias es algo diferente del currículum nacional vigente en cada país, aunque algunos países ya han empezado a reformar sus currículum nacionales para dar por buenas las prescripciones del marco de competencias de PISA.

En materia de estudios internacionales de evaluación de los aprendizajes de los alumnos hay dos aproximaciones distintas y no conciliables. Por una parte, los estudios PISA han insistido en la importancia de medir cómo los sistemas escolares de los diferentes países contribuyen a que sus alumnos adquieran las competencias que el mismo programa PISA considera que son las críticas para el mundo contemporáneo, singularmente en respuesta a las necesidades de desarrollo económico. Pero, por otra parte, otros programas internacionales, como por ejemplo TERCE (UNESCO) en América Latina, TIMSS o PIRLS, así como la mayoría de sistemas nacionales de evaluación, intentan evaluar los logros de los alumnos en relación a las competencias definidas en los correspondientes currículum nacionales. Ni que decir tiene que la comparación internacional de los resultados de los alumnos respecto de su respectivo currículum nacional es extremadamente compleja; así, por ejemplo, hay que encontrar un mínimo común denominador a partir del cual generar el marco de evaluación.

La distinción entre ambas aproximaciones es muy importante. Mientras los sistemas que miden las competencias basadas en el currículum nacional pueden efectivamente predicar algo acerca del comportamiento del sistema escolar en relación a sus objetivos nacionales, los estudios PISA sólo pueden hacerlo con referencia a un currículum teórico que no tiene porque ser necesariamente el que los programas nacionales prescriben y los docentes se esfuerzan a poner en práctica. Por tanto, cuando los resultados de un estudio PISA enjuician la posición relativa de un país en cuanto a las competencias adquiridas por los alumnos en una determinada área curricular lo hacen tomando como criterio de evaluación lo que el Programa juzga que son las competencias básicas. Si la evaluación resultante ubica a un país en una

zona por debajo de la media internacional, la interpretación más simplista asume que esto es indicativo de una baja calidad de la educación. En realidad, esto no tiene porqué ser así, sencillamente porque el sistema puede ser extremadamente eficiente en cuanto a los aprendizajes que el curriculum escolar nacional prescribe, o al revés. Los resultados de un mismo país pueden variar enormemente según cuál sea la aproximación tomada; así, por ejemplo, en TIMSS 1999 Finlandia ocupó el lugar 14 en matemáticas y 10 en ciencias de un total de 38 países participantes. La confusión se produce cuando se asume que lo que miden los estudios PISA y lo que prescribe el curriculum nacional son exactamente lo mismo.

Dos ejemplos ayudarán a aclarar esta cuestión. El primero proviene de Inglaterra, donde el informe nacional de resultados en PISA 2006 incluyó una comparación con los curriculum nacionales en las tres áreas (Bradshaw, Sturman, Vappula, Ager, & Weather, 2007). Las diferencias son importantes en las tres áreas. Por ejemplo, en ciencias el informe sostiene que la prueba PISA coloca mucho mayor énfasis en la información contextual que las pruebas GCSE, que son aquellas bajo cuyo marco se preparan los alumnos ingleses, hasta el punto de que «a pesar de que los efectos no pueden ser cuantificados, algunos alumnos pueden encontrarse con que el estilo y las preguntas de la prueba PISA no les sean familiares. Esto puede afectar sus resultados, especialmente cuando se trata de alumnos que se encuentran en las bandas más bajas de competencias...» (p. 16). Otro tanto se dice del curriculum en matemáticas (p. 17). El segundo ejemplo viene de España¹⁰ y se refiere al área de matemáticas. Al comparar las exigencias del curriculum nacional con el marco de evaluación de PISA, Goñi Zabala (Goñi Zabala, 2011) concluye que, aunque la prueba PISA da más importancia que la habitual en España a los aspectos relacionados con estadística y probabilidad, «el resto de contenidos, aunque nombrados de otra manera, es similar». Sin embargo, «la diferencia está en el nivel... El nivel de la mayoría de los conocimientos a aplicar en estas pruebas puede estimarse que se sitúa sobre 1º de ESO. Una hipótesis plausible es que los rendimientos (de los alumnos españoles) se resienten porque muchos de esos conocimientos¹¹ ya han sido

¹⁰ Y, dicho sea de paso, muy probablemente contradiga las expectativas de los lectores de este país.

¹¹ Nótese, de paso, la insistencia con la que el autor citado insiste en el uso del término «conocimientos» frente a «competencias».

olvidados como consecuencia de un aprendizaje escasamente significativo» (p. 64). De modo semejante, otros autores españoles (Gil Martín & González Aguado, 2011) llegan a sugerir que el ámbito de competencias cubierto por PISA en ciencias es poco menos que incompleto, «porque faltan, por ejemplo, aspectos como; la emisión de hipótesis, el diseño y realización de experimentos, la modernización, el trabajo experimental, el trabajo en equipo...» (p. 67). Queda por ver qué pasaría con los resultados de los alumnos españoles si estas áreas también estuvieran cubiertas.

La peculiar aproximación de los estudios PISA a la evaluación de competencias, tomando como marco de referencia su propia definición curricular, tiene además importantes implicaciones políticas que han sido objeto de críticas en numerosas ocasiones. Fundamentalmente, estas críticas pueden resumirse en dos grandes grupos.

El primer grupo de críticas gira alrededor de la idea de que el Programa PISA prescribe un curriculum internacional al que, en realidad, los países participantes se adhieren, a pesar de que esto pueda llevar a contradicciones con las prescripciones curriculares efectivamente existentes a escala nacional (Fuchs, 2003). En otras palabras, el Programa PISA dicta a los países participantes lo que sus alumnos deberían aprender sin que este dictado haya sido objeto de un debate nacional con los actores relevantes. Es más, en la mayoría de los países esta velada imposición curricular acostumbra a pasar absolutamente inadvertida porque, entre otras razones y como se acaba de indicar, apenas se realizan esfuerzos de comparación entre las exigencias de las pruebas PISA y los curriculum nacionales. Curiosamente, tanto por su impacto mediático como por su penetración en el discurso político, el fenómeno resultante, que generalmente pasa desapercibido, es que la forma en que los estudios PISA miden los logros de los alumnos se convierte, de forma sutil, en un prescriptor de los objetivos de las políticas educativas nacionales al establecer los estándares de competencias de los alumnos de 15 años de edad para todos los países desarrollados y para los que aspiran a serlo. Inadvertidamente, cuando un gobierno acepta sumarse al Programa acepta que el criterio de medida de los logros de aprendizaje de sus alumnos será, no el curriculum nacional, sino el marco de evaluación de competencias establecido por PISA. La asunción que entonces se realiza es muy simple: hay que mejorar la posición relativa del propio país en el concierto internacional porque de ello depende la competitividad económica. Es decir, el criterio de medida se con-

vierte inevitablemente en un objetivo político *per se*, sin que haya mediado necesariamente ninguna discusión entre los actores a escala nacional (Grek, 2009). Es el advenimiento de la gobernanza global de la educación.

Esta crítica no es muy distinta de la que se produce en relación al denominado «gobierno por las cifras» (Grek, 2009) y que se encuentra, de hecho, en todas las esferas de la actividad pública. Es una manera sutil de intervenir políticamente desde la esfera internacional en la nacional sin que se dé la impresión de que se está produciendo semejante intervención. De ahí que se haya acuñado la expresión de «poder suave» para referirse a la influencia de los estudios PISA sobre las políticas nacionales de educación (Bieber & Martens, 2011). Como sucede generalmente con los sistemas de indicadores, y más en general, con las evaluaciones internacionales de políticas públicas y de gobierno (Besançon, 2003), los indicadores elegidos no son nunca neutros sino que siempre reflejan una particular concepción de los objetivos de las políticas públicas (Arndt & Oman, 2006; Kaufmann, Kraay, & Mastruzzi, 2006). Ya lo destacó Nóvoa (Nóvoa, 2010) en las páginas de esta misma revista en relación a la educación en Europa, «la invención de indicadores comparables no es sólo una operación para describir la realidad, sino también una poderosa forma de construir nuevas ideas y prácticas en educación». En suma, quien acepta que le comparen a los demás por su estatura acepta, consciente o inconscientemente, que es algo por lo que hay que preocuparse.

Un segundo grupo de críticas sostiene que el curriculum prescrito por PISA obedece a una particular visión del rol que la enseñanza debe jugar en el desarrollo económico del país (Tröhler, 2010). La definición de las competencias parte de un criterio de significación no tanto para el desarrollo individual del alumno, como para la competitividad y el desarrollo económico del país y el rol que el alumno de hoy deberá desempeñar como trabajador mañana, particularmente en el contexto de la economía del conocimiento¹².

¹² En el camino recorrido desde 1997, cuando la OCDE lanzó el Proyecto DeSeCo (*Definition and Selection of Competencies: Theoretical Foundations*, www.deseco.admin.ch) hasta la definición de las competencias en el marco de evaluación de los estudios PISA se han perdido notables aspectos relacionados, precisamente, con el desarrollo individual, comunitario y social. Como ya señalaron, Bolívar y Pereyra (Bolívar & Pereyra, 2006), «el concepto holístico de competencias que tiene DeSeCo no casa realmente con la estrategia cognitiva y metodológica de PISA».

Como ha señalado reiteradamente el director de los estudios PISA, su propósito consiste en «medir hasta qué punto los jóvenes adultos, a la edad de 15 años y por tanto al aproximarse a la finalización de la escolarización obligatoria, están preparados para enfrentarse a los desafíos de las actuales sociedades del conocimiento -que es lo que PISA expresa cuando se refiere a los alfabetismos¹³. La evaluación está orientada hacia el futuro y se centra en la capacidad de los jóvenes para usar sus conocimientos y competencias para enfrentarse a los desafíos del mundo real, más que en saber hasta qué punto dominan un currículum escolar en particular» (Schleicher, 2006).

PISA, pues, apunta a las necesidades de formación del capital humano para una economía globalizada y las competencias requeridas e interviene así «en la fabricación de seres humanos para una imaginada sociedad global» (Bolivar, 2010). La interrelación entre el concepto de sociedad y economía del conocimiento y competencias es clara: los países compiten para hacerse con una porción mayor del mercado global y necesitan, por consiguiente, un capital humano capacitado para hacerlo. A los jóvenes se les prepara para dar salida a esta necesidad, lo cual ha de permitirles igualmente sobrevivir en ella. La educación se transforma en un bien o mercancía que está, ahora más que nunca, vinculada al sistema económico. Y los estudios PISA tienen la misión de ayudar a los países a comprender cómo su sistema escolar cumple con estas exigencias en un contexto de competencia internacional (Wagemaker, 2008). Y lo hace definiendo las competencias que cuentan, desechando otras como, por ejemplo, las relacionadas con los valores cívicos cuya importancia para la cohesión social y para el individuo¹⁴.

En un informe de evaluación del impacto de los estudios PISA en la política educativa de los países de la OCDE (OECD, 2008) se afirma, abiertamente, que «algunos países se ven ya parte de la economía global y tienden a valorar las competencias evaluadas en PISA... Los países que están conscientemente modernizando su sistema educativo están cambiando sus fines y generalmente hacen uso de los propuestos por PISA. Estos países apuestan por el desarrollo de ciudadanos globales que sean capaces de vivir y trabajar en cualquier lugar del mundo, adoptan para sí las competencias evaluadas

¹³ La expresión original en inglés es *literacy*.

¹⁴ Este es, singularmente, el caso del Proyecto CIVICS de la IEA (Torney-Purta, Lehmann, Oswald, & Schultz, 2001).

en PISA y examinan cómo incorporarlas en sus curriculum» (p. 10). Puede que el puente que los estudios PISA establecen hacia una prescripción curricular universal o, mejor aún, globalizada no sea visible a simple vista, pero algunos países, como por ejemplo Alemania (Kotthoff & Pereyra, 2010; Leschinsky, 2005), Austria (Bozkurt, Brinek, & Retzl, 2007) o los países nórdicos (Kobarg & Prenzel, 2009), ya lo están cruzando.

3.2. El puente más largo: la prescripción de políticas

Como su director ha señalado reiteradamente (Schleicher, 2006), el verdadero objetivo de los estudios PISA es el de ayudar a los países participantes a mejorar sus políticas educativas por medio del análisis de los factores que explican los resultados de los alumnos y la forma en que estos pueden ser objeto de intervención política. Se trata, por lo tanto, de avanzar en el terreno de la formulación de políticas y del diseño de reformas educativas a partir de la evidencia empírica (OECD, 2007a). De este modo, si se cuenta con la evidencia empírica suficiente se puede prescribir qué es lo que funciona, es decir, qué intervención política debe ponerse en práctica porque la evidencia demuestra que tendrá éxito: a la prescripción por la evidencia (Slavin, 2008). Los estudios PISA han intentado establecer el puente desde su propia base de evidencias hasta la prescripción política desde tres perspectivas distintas y sucesivas en el tiempo: la resolución de la función de productividad en educación, la descripción de las políticas educativas de los países que lideran el *ranking* en PISA y, finalmente, el análisis de trayectorias. Las dos primeras opciones han dado lugar a una mejora notable de la base de conocimientos acerca de la educación, pero tienen importantes limitaciones en cuanto a su poder prescriptivo.

En primer lugar, la función de productividad permite explicar los resultados de los alumnos en base a variables independientes que son conocidas, pero no da idea necesariamente de la causalidad subyacente que el observador debe interpretar desde su propio marco teórico de referencia. Hay un salto importante a realizar desde la explicación econométrica hasta la prescripción política. Y este salto puede pasar inadvertido para el lector poco cualificado.

La forma en que los informes de PISA abordan, por ejemplo, la relación entre recursos y resultados, una cuestión extremadamente controvertida (Schneider, 2009), da idea de lo tenue que es la barrera que separa el análisis de la prescripción. En el análisis de resultados en ciencias de PISA 2006 (OECD, 2007b) se exponen una serie de limitaciones metodológicas para prevenir al lector adecuadamente; así, se comentan, entre otros, aspectos relacionados con el carácter no longitudinal de los resultados y se advierte explícitamente de que algunos factores potencialmente importantes no fueron medidos y otros no lo fueron apropiadamente. En concreto, en el informe se lee que «la combinación de estas restricciones limita la capacidad que PISA puede tener de ofrecer estimaciones estadísticas directas de los efectos de los recursos a disposición del centro escolar sobre los resultados educativos» (p. 215). Y, aún a pesar de estas severas limitaciones debidamente hechas explícitas, el informe termina concluyendo con recomendaciones políticas a partir de esta endeble base. Es más, concluye que «lo que es destacable acerca de los efectos más fuertes medidos en este capítulo es que no son los que más estrechamente relacionados están con los recursos materiales finitos, como la distribución de buenos docentes. Más bien, estos efectos se relacionan con la forma en que los centros y el sistema escolar están gestionados; por ejemplo, con la cantidad de tiempo que los alumnos pasan en el aula y hasta qué punto las escuelas tienen que rendir cuentas de sus resultados» (p. 277). Dicho con otras palabras, si los centros escolares rinden cuentas todo lo demás, incluyendo la provisión de buenos docentes, tendrá un efecto marginal despreciable. Pero, ¿cuál es la evidencia empírica para la rendición de cuentas, es decir, para que los centros escolares observen mejoras si se publican los resultados de sus alumnos? A decir verdad, el propio informe ofrece datos que contradicen esta evidencia: en los cinco países que en esa oleada de PISA obtuvieron mejores resultados y además con un menor impacto de la extracción socioeconómica de los alumnos¹⁵ (Australia, Canadá, Finlandia, Japón y Corea) el porcentaje de alumnos que asiste a escuelas que publican sus resultados va del 4% (Finlandia) al 64% (Canadá) y la media de estos cinco países es del 31%, cuando la media del conjunto de países de la OCDE es del 38%. En suma, las prácticas de los mejores países sugerirían precisamente que lo que hay que hacer es reducir el porcentaje de alumnos

¹⁵ En este volumen ésta es la acepción que se da de la rendición de cuentas.

que van a escuelas que publican sus resultados y, por consiguiente, reducir la presión en este sentido sería la política recomendable. ¿Será entonces que la agenda política, y no la evidencia, es la que prima la dirección de la prescripción sugerida?

En segundo lugar, la descripción de las políticas seguidas por los países que obtienen mejores resultados puede ser enormemente informativa y una ventana de oportunidad para ensanchar el horizonte de las opciones políticas en materia educativa, algo que ha explorado tanto la OCDE, a petición de Alemania (OECD, 2004), como otros, como el conocido informe McKinsey, con la abierta ambición de explicar «cómo los mejores sistemas escolares del mundo llegaron a la cima» (Barber & Mourshed, 2007). Pero, esta función informativa dista mucho de tener un carácter predictivo acerca de cuáles de esas políticas que caracterizan a los países con mejores resultados funcionarán eventualmente en otros contextos y llevarán a los demás países a mejorar los logros de sus alumnos. En realidad, el problema de esta segunda aproximación es que cae en la típica trampa de seleccionar los casos en función de la variable dependiente. De este modo, las políticas y las prácticas de los países con mejores resultados se presentan como patrones a seguir sin que medie ninguna evidencia acerca de sus eventuales efectos positivos, a medio o a largo plazo, en países con resultados mediocres.

No hay ninguna presentación de estrategias alternativas ni ninguna discusión acerca de la importancia del contexto, ni mucho menos se ponen de manifiesto las eventuales contradicciones que pueden hallarse en un mismo sistema. Por ejemplo, el citado informe McKinsey sitúa a Corea como un ejemplo a seguir en materia de políticas de atracción de candidatos con un perfil académico alto a la carrera docente en la enseñanza primaria, pero señala igualmente que éste no es el caso de la enseñanza secundaria, donde el país se encuentra con que el «estatuto y el atractivo de la profesión han disminuido en Corea del Sur, convirtiéndola en poco atractiva a los ojos de candidatos con buenos resultados académicos». En los resultados de PISA 2006, Corea se situó entre los mejores y, por cierto, con uno de los mayores porcentajes de jóvenes en el nivel más alto de desempeño en matemáticas. Entonces, ¿la buena política consiste en hacer atractiva la profesión docente en primaria y olvidarse de la enseñanza secundaria? ¿O es que, en esto, hay que seguir mejor el modelo finlandés?

La tercera perspectiva, centrada en el análisis de trayectorias, que es la que ahora se ensaya con mayor ahínco (OECD, 2010b), merece un análisis más detallado. Por trayectorias educativas cabe entender la evolución de los resultados de los alumnos en una serie temporal suficientemente larga como para permitir apreciar el impacto de determinadas políticas o reformas. El análisis comparativo internacional de estas trayectorias debería permitir identificar qué países han conseguido mejorar los logros de sus alumnos, o los han empeorado, y, en segundo lugar, estudiar qué políticas públicas pueden explicar la correspondiente trayectoria con la finalidad, eventualmente, de predicar su generalización a otros países en circunstancias semejantes. Esto es lo que el segundo informe McKinsey pretende hacer bajo el sugerente título de «cómo se convierte un sistema educativo de bajo desempeño en uno bueno»¹⁶ (Mourshed, Chijioko, & Barber, 2010).

El programa PISA no empezó las primeras evaluaciones hasta el año 2000 e, incluso ahora con cuatro oleadas ya completadas, el número de países que han experimentado mejoras significativas es muy reducido¹⁷. Esto puede interpretarse de tres maneras diferentes:

- a) la mayoría de países no han hecho nada para mejorar los resultados de los alumnos, es decir, no han puesto en práctica políticas públicas apropiadas;
- b) han hecho lo que han podido, pero sin éxito; y
- c) diez años (2000-2009) no es un lapso de tiempo suficiente para que las reformas escolares lleguen a tener un reflejo en los resultados de aprendizaje de los alumnos.

Así las cosas, no es extraño que los análisis comparativos internacionales más conocidos hasta el momento, lejos de centrarse en el análisis de las trayectorias (atendido el reducido número de países que han variado de manera significativa su posición relativa en las comparaciones internacionales de resultados), prefieran más bien analizar qué políticas han puesto en práctica

¹⁶ Este es el título del resumen ejecutivo del informe en español, pero la traducción literal del título original en inglés sería la siguiente: «¿Cómo los sistemas educativos que más mejoran continúan mejorando?»

¹⁷ Hasta la publicación de los resultados de PISA 2009 el único ejemplo de trayectoria positiva en PISA ha venido siendo Polonia (Jakubowski, Patrinos, Porta, & Wisniewski, 2010).

los países que sistemáticamente obtienen buenos resultados. Los análisis propuestos, por lo tanto, se centran en ver cuáles son las concomitancias y los elementos comunes entre estos países, cosa que no deja de ser muy interesante, pero lo cierto es que el fundamento empírico y la aproximación metodológica de estos análisis puede ser fácilmente contestado. Dicho de otra manera, no existe un vínculo de causalidad demostrable y, además, teniendo en cuenta la disparidad de contextos es difícil pensar que las mismas estrategias puedan servir igualmente a las necesidades particulares de todos y cada uno de los países. Hay que ser, pues, muy prudente con la confianza que se deposita en estos análisis.

Será, en definitiva, un puente largo y arduo de cruzar. Pero, suponiendo que se consiguiera, quedaría pendiente una importante cuestión que pone siempre a la picota a las reformas basadas en la evidencia. Es la cuestión de la legitimidad de las políticas, que para bien o para mal no proviene de su base empírica sino del grado de aceptación política y pública de la medida propuesta y que es algo que tiene mucho que ver con el contexto y los posicionamientos estratégicos de los actores implicados (Waldow, 2009). Una buena reforma destinada a implantar la comprensividad en la enseñanza secundaria, por ejemplo, puede fracasar, lisa y llanamente, si los cambios que se operan en el contrato docente (salarios y condiciones laborales) no compensan del esfuerzo adicional requerido. La lección es que cuando cualquier intento de racionalizar la política o la administración educativa ignora aspectos importantes de la lógica del funcionamiento del sistema tiene grandes posibilidades de fracasar. Y uno de los aspectos que con más frecuencia se ignoran es que los actores políticos enjuician las propuestas políticas no sólo desde la perspectiva de la solución de problemas sino, tanto o más si cabe, desde la perspectiva del refuerzo de su propia legitimidad en tanto que actores y de su acción política.

4. LA AGENDA PENDIENTE DE INVESTIGACIÓN

Los análisis anteriores plantean implícitamente la agenda pendiente de investigación en relación a la ambición prescriptiva en materia política de los estudios PISA. Esta agenda debería contener, fundamentalmente, tres elementos: el análisis de trayectorias, la observación de las prácticas en el aula,

y, finalmente, la triangulación con otros estudios empíricos, tanto a escala internacional como nacional.

4.1. El análisis de trayectorias

El gran objetivo de los estudios PISA ha sido desde el principio informar las políticas públicas de educación proporcionando evidencias empíricas acerca de cuáles son las que permiten que los sistemas escolares aumenten el nivel de logros de los alumnos. En un ámbito como el de las políticas de educación, donde apenas existe una cultura de la evidencia empírica, esta ambición sigue siendo muy necesaria para informar tanto los debates públicos como los procesos de toma de decisiones.

La agenda de investigación está, pues, clara pero el enfoque seguido hasta el momento tiene que mejorar sustancialmente. Por una parte, el énfasis inicial centrado en la pregunta cuáles son las convergencias políticas entre los sistemas que mejores resultados obtienen no responde necesariamente a la pregunta de qué es lo que hay que hacer para mejorar. Como se ha hecho notar con frecuencia, las razones que explican la preeminencia de ciertos países, como Finlandia o Corea, van mucho más allá de la puesta en práctica de políticas públicas en educación y tienen mucho que ver con el texto político, económico y social más amplio de estos países; algo que, por supuesto, ningún otro país puede reproducir. Por otra parte, la segunda vuelta de tuerca que los estudios PISA han dado a esta cuestión preguntándose acerca de cuáles son los países cuya trayectoria ha mejorado significativamente a lo largo del tiempo tampoco ha dado un resultado a la altura de las expectativas. Como ya se ha indicado, el número de países que caen en esta categoría no sólo es extremadamente reducido sino que, de nuevo, presentan contextos muy dispares. Además, una vez identificados los casos el verdadero problema metodológico es cómo establecer una causalidad retroactiva, es decir, cómo demostrar empíricamente que la mejora en los resultados de PISA está significativamente vinculada a unas determinadas políticas. Los estudios de caso pueden ayudar pero inevitablemente pondrán el acento en el contexto de las políticas igualmente

La aproximación requerida debería ser algo menos ambiciosa y, al mismo tiempo, más fiel a la realidad del mundo de las políticas educativas. La pregunta de investigación verdaderamente abordable es simple: ¿qué políticas

específicamente diseñadas para mejorar los logros de los alumnos, tal y como los miden los estudios PISA, han tenido éxito? La pregunta remite de nuevo a estudios de caso pero partiendo del principio de que el universo de las políticas públicas potencialmente significativas para la mejora de los resultados de los aprendizajes es más reducido y, muy probablemente, mucho más relacionado con lo que ocurre dentro del aula, algo todavía ausente en la ecuación explicativa resultante de los estudios PISA.

4.2. La inclusión del aula en el modelo explicativo

Las paredes del aula siguen marcando una gran frontera que hasta el momento los estudios PISA no han sido capaz de franquear. No es extraño que sea así puesto que, en el marco metodológico de PISA, hay dos cuestiones difícilmente resolubles:

- ¿Sería posible tomar una muestra representativa de las prácticas de enseñanza en cada área de conocimientos (lengua, matemáticas, ciencias) en cada país, para permitir una comparación internacional? Y, si así fuera, ¿sería posible hacerlo de forma que la varianza estadística quedara igualmente reflejada para dar cuenta así de los distintos modelos de práctica pedagógica que pueden convivir incluso dentro de un mismo centro escolar?
- ¿Sería posible vincular patrones de prácticas docentes con logros de los alumnos? Dicho de otra manera, como los estudios PISA miden los logros acumulados por los alumnos a lo largo de toda su escolaridad, la cuestión es esbozar qué patrones, esto es, combinaciones sucesivas de prácticas docentes, conducen a mejores logros.

Las dos cuestiones, que podrían formularse fácilmente en términos de preguntas de investigación, no son fácilmente resolubles. Los dos problemas metodológicos fundamentales son cómo observar y documentar las prácticas docentes y, en segundo lugar, cómo establecer patrones de prácticas que abracen en realidad todo el período significativo que, quizás para empezar, podría circunscribirse meramente a la enseñanza secundaria obligatoria.

El horizonte de investigación es extremadamente amplio. Y no estará de más decir que implicará la entrada en juego de métodos de investigación empíricos que, a la fuerza, incorporarán la observación directa del aula cuyos

problemas en términos de validez y confianza son sobradamente conocidos. Pero probablemente no haya otro camino.

En realidad, no estará de más señalar que si no directamente el Programa PISA, por lo menos TALIS sí tiene previsto abordar por medio de grabaciones de sesiones de clase la cuestión de las prácticas docentes. Lejos de buscar un objetivo ambicioso, en un primer estadio se trata exclusiva de documentar prácticas, no necesariamente con carácter representativo. Del mismo modo, para avanzar igualmente en la vinculación entre prácticas docentes y resultados de los alumnos, TALIS pretende igualmente introducirse más a fondo en la cuestión de cómo las prácticas docentes (a partir de las declaraciones de los propios docentes) pueden de una u otra forma tener incidencia en los logros de los alumnos. Pero está por ver hasta dónde se podrá llegar.

4.3. La triangulación

Otro aspecto importante en la agenda es la triangulación de los estudios PISA con otros estudios de ámbito internacional (como TIMSS o TERCE, por ejemplo), por una parte, así como con evaluaciones de carácter nacional, por otra, algo que lógicamente compete a los gobiernos nacionales y a los investigadores ([Hamilton, 2003](#)). La idea es muy simple: se trata de ver si las variaciones que un país obtiene en las sucesivas oleadas de PISA concuerda con las que se observan en otros estudios comparativos internacionales de carácter empírico y con los ejercicios nacionales de evaluación.

A simple vista, se podría pensar que esto no podría ser de otra manera pero un examen más detallado de la cuestión plantea no pocos problemas, dejando de lado detalles importantes estrictamente organizativos o de diseño (como la frecuencia de las oleadas o la edad de los alumnos a quienes se examina) ([Rindermann, 2007](#)). El primero y más importante de estos problemas es el de la alineación de los dos estudios; dicho de otra manera, no está claro que los restantes estudios internacionales o los estudios nacionales de evaluación den cuenta exactamente de lo mismo que se evalúa en los estudios PISA. Así, por ejemplo, los estudios nacionales pueden primar la evaluación de los logros de los alumnos en relación al currículum nacional y, no en cambio, en relación al marco de competencias de PISA.

Pese a todo, esta triangulación resulta del todo punto imprescindible para tener una perspectiva comprensiva acerca de los logros de los alumnos. De no hacerlo así, se caería en el error de considerar a los estudios PISA como la única fuente de evidencia empírica para la formulación y la evaluación de políticas públicas en educación.

5. CONCLUSIÓN, CON UNA MIRADA HACIA EL FUTURO DE LOS ESTUDIOS COMPARATIVOS DE EDUCACIÓN

Nada ha contribuido a cambiar más la forma en que tanto la opinión pública como el discurso político se aproximan a la educación que los estudios PISA. Del mismo modo, nada ha contribuido de forma más decisiva a la expansión de la cultura de la evaluación en el sector educativo, hasta el punto de resultar en una verdadera explosión de centros e institutos de evaluación educativa prácticamente en todo el mundo, con la doble intención de coordinar la participación nacional en evaluaciones internacionales, así como diseminar sus resultados, y de generar programas nacionales de evaluación de los logros de los alumnos.

Pero los estudios PISA no sólo arrojan una mirada comparativa internacional acerca de los logros de los sistemas educativos sin más. Van mucho más allá en su carácter prescriptivo para las políticas educativas a escala nacional. Y lo hacen de creando puentes desde el análisis de los resultados hacia la prescripción política que toman dos formas complementarias: la creación de estándares de facto para todos los países y prescribiendo abiertamente qué políticas públicas de educación pueden llevar a todos los países por igual a mejorar sus resultados por medio de reformas.

En primer lugar, los estudios PISA han generado un marco de referencia universal que define las competencias básicas que los alumnos de 15 años deberían adquirir en las áreas de lenguaje, matemáticas y ciencias. Al hacerlo, han prescrito de forma sutil lo que los currículum nacionales deberían contener, dejando espacio para la nota de color que cada país quiera añadir en términos de otras áreas curriculares. En definitiva, cualquier país que acepta que sus alumnos sean evaluados por el Programa PISA acepta abiertamente que se trata de una buena medida de los logros escolares. Y, sin embargo, como se ha mostrado en esta contribución es más que posible que el

trabajo de los docentes en el aula, lo mismo que las estructuras y las regulaciones nacionales, obedezca a otra lógica distinta que cabría esperar mucho mejor reflejada en evaluaciones organizadas por el propio país. Cuando los estudios PISA acreditan, por ejemplo, la posición mediocre de determinado país en el *ranking* internacional se infiere automáticamente que su sistema educativo no funciona bien y que, por consiguiente, tiene que ser reformado. Pero es perfectamente plausible que el sistema funcione perfectamente bien de acuerdo a los dictados internos; lo que no hace bien es seguir los del Programa PISA.

Dejando de lado las frecuentes críticas que el Programa PISA, y la OCDE, han recibido por intentar crear un curriculum global en aras de mejorar el crecimiento económico, los países deben aprender a relativizar los resultados, cosa que pueden hacer de tres formas. Primero, elevando las competencias técnicas de quienes, como políticos o administradores de la educación, deben responder a la pregunta de cuál es la calidad real del sistema educativo de forma que entiendan bien qué hace PISA y cuáles son sus límites, singularmente cuando extiende puentes hacia la prescripción política. Segundo, triangulando los estudios PISA con otros estudios internacionales (como TIMSS o TERCE) y con las evaluaciones nacionales, lo cual exige una inversión importante en investigación sobre educación. Tercero, elevando el tono crítico y la exigencia con respecto a la forma en que el Programa PISA genera sus recetas de reforma política, infiriendo causalidades allí donde tal vez no las haya.

Al margen de estas consideraciones hay otro aspecto no menos importante, por lo menos en el contexto de esta publicación y de su audiencia. Los estudios PISA, como sus equivalentes, atraen a académicos y técnicos que raramente proceden de las facultades de educación y pocos son los expertos en educación comparada que llegan más allá de un análisis centrado en el marco conceptual o en los usos políticos. Probablemente haya más publicaciones utilizando datos de PISA en las revistas de economía, de sociología, de ciencia política o de psicometría que en las de educación, y no sólo en nuestro país.

Sea como sea, uno de los grandes beneficiarios de la creciente relevancia pública y política de los estudios PISA es la investigación acerca de la educación y su diálogo con los procesos políticos. Pero este diálogo es, en rea-

lidad, un campo de batalla donde se juega desde dos bandos: por una parte, los actores políticos que buscan soluciones *quick and dirty* que aporten algo en relación a su agenda; y, por otra, los investigadores cuya pauta y ritmo de trabajo obedece a la lógica, tranquila y transparente, de la ciencia. ¿Será la investigación sobre educación capaz de satisfacer las expectativas, a veces poco realistas, de los políticos y de la opinión pública en materia de resolución de problemas?

Pero la investigación acerca de la educación no es sólo patrimonio de las facultades de educación. Es más, las facultades de educación y en particular los programas de formación en educación comparada tienen un reto pendiente: el de cualificar adecuadamente a los futuros titulados para que sean capaces no sólo de tener una aproximación crítica a estudios como PISA sino también de manejar apropiadamente los recursos, datos y análisis que ofrecen. No hacerlo significa sustraer oportunidades a los profesionales de la educación y, en particular, a los que desean desarrollar una carrera en el ámbito de los estudios comparados o de la política educativa.

Como se ha dicho y repetido hasta la saciedad, PISA ha venido para quedarse. Y el espacio de relevancia que está encontrando tanto en las discusiones expertas como en la opinión pública estaría mucho mejor atendido si también la voz de aquellos que se forman en las facultades de educación se hiciera presente de forma técnicamente cualificada.

REFERENCIAS BIBLIOGRÁFICAS

- ANDERSEN, F. O. (2010). Danish and Finnish PISA results in a comparative, qualitative perspective: How can the stable and distinct differences between the Danish and Finnish PISA results be explained? *Educational Assessment Evaluation and Accountability*, 22(2), 159-175.
- ARNDT, C., & OMAN, C. (2006). *Uses and Abuses of Governance Indicators*. Paris: OECD Publishing.
- BAKER, D. P., & LETENDRE, G. K. (2005). *National Differences, Global Similarities. World Culture and the Future of Schooling*. Stanford: Stanford University Press.
- BARBER, M., & MOURSHED, M. (2007). *How the world's best performing school systems come out on top*. London: McKinsey and Company.

- BESANÇON, M. (2003). *Good Governance Rankings: The Art of Measurement*. Cambridge: World Peace Foundation.
- BIEBER, T., & MARTENS, K. (2011). The OECD PISA Study as a Soft Power in Education? Lessons from Switzerland and the US. *European Journal of Education*, 46(1), 101-116.
- BOLIVAR, A. (2010). PISA como discurso: el espacio público educativo. *Organización y gestión educativa*, 18(86), 18-21.
- BOLIVAR, A., & PEREYRA, M. (2006). El proyecto DeSeCo sobre definición y selección de competencias clave. Prólogo a la edición española. In D. S. Rychen & L. H. Salganik (Eds.), *Las competencias clave para el bienestar personal, social y económico* (pp. 1-33). Málaga: Aljibe.
- BOZKURT, D., BRINEK, G., & RETZL, M. (2007). PISA in Österreich: Mediale Reaktionen, öffentliche Bewertungen und politische Konsequenzen. In S. T. Hopmann, G. Brinek & M. Retzl (Eds.), *PISA zufolge PISA — PISA according to PISA*. Wien: Lit Verlag.
- BRADSHAW, J., STURMAN, L., VAPPULA, H., AGER, R., & WEATHER, R. (2007). *Achievement of 15-year-olds in England: PISA 2006 National Report*.
- BRINT, S. G. (2006). *Schools and Societies* (2 ed.). Stanford: Stanford University Press.
- BRUNNER, M., ARTELT, C., KRAUSS, S., & BAUMERT, J. (2007). Coaching for the PISA test. *Learning and Instruction*, 17(2), 111-122.
- CHROMY, J. R. (2002). Sampling issues in design, conduct, and interpretation of international comparative studies of school achievement. In A. C. Porter & A. Gamoram (Eds.), *Methodological advances in cross-national surveys of educational achievement* (pp. 80-114). Washington DC: National Academy Press.
- DALL, A. (2010). *PISA and reading literacy — comparing apples and oranges?* Paper presented at the 23rd International Congress for School Effectiveness and Improvement (ICSEI).
- DANIELSEN, A. G., SAMDAL, O., HETLAND, J., & WOLD, B. (2009). School-related social support and students' perceived life satisfaction. *Journal of Educational Research*, 102(4), 303-318.
- DIXON, R., MULLER, S. M., ARNDT, C., & HOOD, C. (2008). Where International Rankings Meet Neativity Bias. Domestic Press Responses to International Educational Rankings in Three EU Countries. Retrieved 25.8.2011, from <http://www.christopherhood.net>

- FUCHS, H. W. (2003). Towards world curriculum? — The concept of basic education (literacy) underlying PISA and the tasks allocated to schooling. *Zeitschrift Fur Padagogik*, 49(2), 161-179.
- GIL MARTÍN, A., & GONZÁLEZ AGUADO, M. E. (2011). Competencia científica. *Cuadernos de Pedagogía*(413), 67-70.
- GOÑI ZABALA, J. M. (2011). La competencia matemática. PISA y el sistema solar: encuentros en la segunda fase. *Cuadernos de Pedagogía*(413), 62-66.
- GREK, S. (2009). Governing by numbers: the PISA 'effect' in Europe. *Journal of Education Policy*, 24(1), 23-37.
- GRUBER, K.-H. (2006). The German "PISA-shock": some aspects of the extraordinary impact of the OECD's PISA study on the German education system. In H. Ertl (Ed.), *Cross-national Attraction in Education: accounts from England and Germany*. Oxford: Symposium Books.
- HAMILTON, L. (2003). Assessment as a Policy Tool. *Review of Research in Education*, 27(1), 25-68.
- HANUSEK, E. A. (1979). Conceptual and empirical issues in the estimation of educational production functions. *Journal of Human Resources*(14), 351-388.
- HANUSEK, E. A. (2003). The failure of input based schooling policies. *The Economic Journal*(113), 64-98.
- HOPMANN, S. T., BRINEK, G., & RETZL, M. (Eds.). (2007). *PISA zufolge PISA — PISA according to PISA*. Wien: Lit Verlag.
- JAKUBOWSKI, M., PATRINOS, H. A., PORTA, E. E., & WISNIEWSKI, J. (2010). *The Impact of the 1999 Education Reform in Poland*. Washington, DC: The World Bank.
- JANIK, T., & SEIDEL, T. (Eds.). (2009). *The Power of Video Studies in Investigating Teaching and Learning in the Classroom*. Dortmund: Waxmann Verlag.
- KAUFMANN, D., KRAAY, A., & MASTRUZZI, M. (2006). *Governance Matters V: Aggregate and Individual Governance Indicators for 1996-2005*. Washington, DC: World Bank Policy Research Working Paper No. 4012.
- KENNEY, M. M. (2005). *Inside Teaching. How Classroom Life Undermines Reform*. Cambridge: Harvard University Press.
- KOBARG, M., & PRENZEL, M. (2009). Keyword: The myth of the Nordic educational systems. *Zeitschrift Fur Erziehungswissenschaft*, 12(4), 597-615.

- KOLESNIKOV-JESSOP, S. (2010, 1 June). The Writing on the Wall. *Newsweek*.
- KOTTHOFF, H. G., & PEREYRA, M. (2010). La experiencia de PISA en Alemania. *Profesorado*, 13(2).
- LESCHINSKY, A. (2005). From the Educational Council to PISA — A study on the recent history of German educational policy. *Zeitschrift Fur Pädagogik*, 51(6), 818-839.
- LEVIN, H. M. (1974). Measuring Efficiency in educational production. *Public Finance Quarterly*(2), 3-24.
- MARTÍNEZ ARIAS, R. (2006). La metodología de los estudios PISA. *Revista de Educación*(número extraordinario 2006), 111-129.
- MOURSHED, M., CHIJOKE, C., & BARBER, M. (2010). *How the world's most improved school systems keep getting better*. London: McKinsey and Company.
- NÓVOA, A. (2010). La construcción de un espacio educativo europeo: gobernando a través de los datos y la comparación. *Revista Española de Educación Comparada*(16), 23-41.
- O'NEIL, H. F., ABEDI, J., MIYOSHI, J., & MASTERGEORGE, A. (2005). Monetary incentives for low-stakes tests. *Educational Assessment Evaluation and Accountability*(19), 185-208.
- OECD. (2004). *What Makes School Systems Perform? Seeing School Systems through the Prism of PISA*. Paris: OECD Publishing.
- OECD. (2007a). *Evidence in Education: Linking Research and Policy*. Paris: OECD Publishing.
- OECD. (2007b). *PISA 2006 Science Competencies for Tomorrow's World*. Paris: OECD Publishing.
- OECD. (2008). *External Evaluation of the Policy Impact of PISA. EDU/PISA/GB(2008)35/REV1*. Paris: OECD.
- OECD. (2009). *Creating Effective Teaching and Learning Environments: First Results from TALIS*. Paris: OECD Publishing.
- OECD. (2010a). *Making Reform Happen . Lessons from OECD Countries*. Paris: OECD Publishing.
- OECD. (2010b). *PISA 2009 Results: What Makes a School Successful? Resources, policies and practices*. Paris: OECD Publishing.
- OECD. (2011). *Building a High-Quality Teaching Profession. Lessons from around the world*. Paris: OECD Publishing.

- OSBORN, M., BROADFOOT, P., MCNESS, E., PLANEL, C., RAVN, B., & TRIGGS, P. (2003). *A World of Difference? Comparing Learners Across Europe*. Maidenhead: Open University Press.
- PRAIS, S. J. (2003). Cautions on OECD'S recent educational survey (PISA). *Oxford Review of Education*, 29(2), 139-163.
- RINDERMAN, H. (2007). The g-factor of international cognitive ability comparisons: The homogeneity of results in PISA, TIMSS, PIRLS and IQ-tests across nations. *European Journal of Personality*, 21(5), 667-706.
- RUDDOCK, G., T., C.-M., PURPLE, C., & AGER, R. (2006). *Validation Study of the PISA 2000, PISA 2003 and TIMSS-2003 International Studies of Pupil Attainment*.
- SCHLEICHER, A. (2006). Fundamentos y cuestiones políticas subyacentes al desarrollo de PISA. *Revista de Educación, Número extraordinario dedicado a PISA*, 21-43.
- SCHNEIDER, M. (2009). The International PISA Test. A risky investment for states. *Education Next*, 9(4), 69-74.
- SLAVIN, R. E. (2008). Evidence-based reform in education: What will it take? *European Educational Research Journal*, 7(1), 124-128.
- SMITH, L. (2008). *Schools that change: Evidence-based improvement and effective change leadership*. Thousand Oaks: Corwin Press.
- STACK, M. (2006). Testing, Testing, Read All About It: Canadian Press Coverage of the PISA Results. *Canadian Journal of Education*, 29(1), 49-69.
- TORNEY-PURTA, J., LEHMANN, R., OSWALD, H., & SCHULTZ, W. (2001). *Citizenship and education in twenty-eight countries: civic knowledge and engagement at age fourteen*. Amsterdam: International Association for the Evaluation of Educational Achievement.
- TRÖHLER, D. (2010). Harmonizing the Educational Globe. World Polity, Cultural Features, and the Challenges to Educational Research. *Studies in Philosophy and Education*, 29(1), 5-17.
- WAGEMAKER, H. (2008). Choices and trade-offs. *Assessment in Education: Principles, Policy & Practice*, 15(3), 267-278.
- WALDOW, F. (2009). What PISA Did and Did Not Do: Germany after the «PISA-sock». *European Educational Research Journal*, 8(3), 476-483.
- WISE, S. L., & DEMARS, C. E. (2003). *Examinee motivation in low-stakes assessment: Problems and potential solutions*. Paper presented at the Annual Meeting of the American Association of Higher Education Assessment Conference.

PROFESIOGRAFÍA

Francesc Pedró

Es doctor en Educación Comparada por la UNED y actualmente es Director de la sección de Evaluación de Políticas Educativas de la UNESCO. Anteriormente fue analista *senior* en el Centro de Investigación e Innovación Educativa (CERI) de la OCDE, donde fue responsable, entre otros, del *Proyecto de Innovación Educativa*. Con anterioridad fue Director Académico del Programa de Calidad Educativa de la Universidad *Pompeu Fabra* y profesor de Educación Comparada y Políticas Públicas de la misma universidad.

Datos de contacto: f.pedro@unesco.org

Fecha de recepción: 29 de septiembre de 2011

Fechas de revisión: 23 de noviembre de 2011 y 23 de noviembre de 2011

Fecha de aceptación: 24 de noviembre de 2011