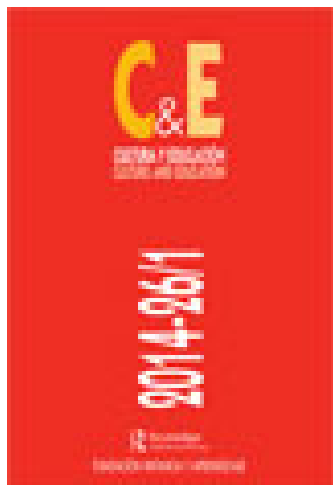


This article was downloaded by: [UGR-BTCA Gral Universitaria]

On: 10 November 2014, At: 09:24

Publisher: Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Cultura y Educación: Culture and Education

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/rcye20>

Luces y sombras de PISA. Sentido educativo de las evaluaciones externas

Angel I. Pérez^a & Encarnación Soto^a

^a Universidad de Málaga

Published online: 23 Jan 2014.

To cite this article: Angel I. Pérez & Encarnación Soto (2011) Luces y sombras de PISA. Sentido educativo de las evaluaciones externas, *Cultura y Educación: Culture and Education*, 23:2, 171-182, DOI: [10.1174/113564011795944758](https://doi.org/10.1174/113564011795944758)

To link to this article: <http://dx.doi.org/10.1174/113564011795944758>

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at <http://www.tandfonline.com/page/terms-and-conditions>

Luces y sombras de PISA. Sentido educativo de las evaluaciones externas

ANGEL I. PÉREZ Y ENCARNACIÓN SOTO

Universidad de Málaga



Resumen

Analizar la función, sentido y consecuencias de las evaluaciones externas de los sistemas educativos basadas en pruebas estandarizadas o test, de carácter internacional, así como sus posibles consecuencias en las políticas nacionales se convierte en el eje de este artículo. Concretamente, nos proponemos identificar las fortalezas y debilidades de PISA, una de las evaluaciones externas que ha marcado en la última década no sólo el ritmo y las iniciativas de políticas educativas sino -y sobre todo- de la mayoría de las informaciones y titulares difundidos sobre educación. Titulares y noticias que han contribuido a crear un superficial marco o estado de la cuestión sobre la calidad de los sistemas educativos, en definitiva un estado de opinión y en consecuencia de decisión y análisis que no ha gozado ni del rigor ni de la independencia, ni de la profundidad requerida.

Palabras clave: Rendición de cuentas, evaluación de los estudiantes, PISA, evaluación sumativa y formativa, competencias clave, evaluación basada en competencias, evaluación auténtica.

PISA's lights and shadows. Educational sense of external assessment

Abstract

The core of this paper is to analyse the function, sense and consequences of international external assessments of educational systems based on standardised tests, and the possible consequences for national policies. Specifically, we aim to identify the strengths and weaknesses of PISA, an external assessment that in the last decade has marked not only the rhythm and initiatives of educational policies, but above all of most spread headlines and news about education. These headlines and news have contributed to creating a superficial framework or state of information on the quality of educational systems; in short, a state of opinion and, as a result, of decision and analysis that do not enjoy the necessary rigour, independence, and depth.

Keywords: Accountability, pupil assessment, PISA, summative and formative evaluation, key competences, competence based evaluation, authentic assessment.

Incremento de las evaluaciones internacionales y políticas de rendición de cuentas: de los controles de entrada a los controles de salida

En el pasado las políticas educativas se basaban en las creencias y opiniones de los líderes hoy pretenden basarse en los hechos: la eficacia de las escuelas y el rendimiento de los estudiantes (Orivel, 2005, pp. 11-12).

El creciente interés en evaluar la calidad de los resultados de la educación se puede ver en el incremento de programas de evaluación internacional como por ejemplo el PIRLS (Progress in Reading and Literacy Study), TIMSS (Third International Mathematics and Science Study) o PISA (International Program of Student Assessment). En todos estos programas se asume que la evaluación y comparación de los sistemas escolares mediante indicadores comunes responde al propósito de favorecer el control social, informado, de la calidad de la educación en servicio público, pues constituyen instrumentos válidos para el diagnóstico y, en su caso, la mejora de la calidad del sistema. Se supone que la comparación de los resultados puede ayudar a detectar problemas, fortalezas y debilidades y a identificar y difundir las buenas prácticas escolares que producen resultados satisfactorios.

Ahora bien, el control social de los servicios públicos en general y de la educación en particular requiere, para ser legítimo, que se cumplan al menos las siguientes condiciones democráticas:

- La participación de los interesados en el análisis y control de las finalidades y sentido del programa de evaluación así como de la coherencia y calidad de cada una de las fases del proceso: diseño, aplicación, corrección, interpretación y difusión de resultados.

- La evaluación de todos los agentes e instituciones que intervienen en el sistema, y por tanto responsables de su funcionamiento, entre los cuales la propia administración, sus políticas y sus modos de gestión ocupan un lugar privilegiado. Evaluar a los docentes sin evaluar la administración, ofrece muy poca credibilidad.

Por otra parte, proponer y utilizar la evaluación externa como instrumento principal de diagnóstico y como detonante privilegiado para provocar el cambio y la mejora, suscita la razonable sospecha de que se inicia la casa por el tejado. ¿Qué motivos hay para proponer o imponer modelos de enseñanza que se consideran valiosos y eficaces, mediante las pruebas externas de evaluación, la puerta de salida, y no desde el diseño-desarrollo del currículum y la formación de los docentes, la puerta principal, de entrada?

Por otra parte, como destaca Linn (2000), el énfasis en las políticas de rendición de cuentas a través de pruebas estandarizadas o tests nacionales o internacionales se debe a un conjunto complejo de factores espurios entre los que cabe destacar: el relativo reducido coste económico, el hecho de que se puedan gestionar y controlar desde fuera del sistema educativo, y la relativa velocidad con que pueden trasladarse los resultados a las estrategias políticas concretas que pretenden la mejora del sistema. Ahora bien, la utilización de estos instrumentos y programas no es inocuo desde el punto de vista pedagógico, y conlleva consecuencias que es necesario indagar y valorar, entre las que destacamos las siguientes:

La rendición de cuentas y el control social parecen sustentarse en la identificación de resultados en términos de indicadores de calidad, y como afirma Normand (2003, 2004), estos indicadores se han transformado, con demasiada frecuencia, de herramientas cognitivas que pueden informar las decisiones políticas a simples herramientas técnicas que permiten que tales procesos sean modificados y legitimados. Con demasiada frecuencia la función informativa se transforma en función normativa.

Por otra parte, el lenguaje que se utiliza en la comunicación de los resultados de las pruebas se ha vuelto tan especializado que las evaluaciones escapan al control crítico de quienes las sufren. Además, como destaca acertadamente Hammerley (2005) las deci-

siones que pretenden basarse en las evidencias, en los indicadores objetivos de calidad, olvidan que los valores desbordan el estado de los hechos y de las evidencias, para situarse en el terreno del debate ideológico, ético y político.

Por otra parte, las exigencias metodológicas y estadísticas de trabajar con variables discretas y no con procesos complejos, induce fácilmente a una concepción también simplista de la vida escolar y de los procesos de enseñanza-aprendizaje como entidades aisladas, discretas e independientes cuyo funcionamiento se puede conocer e incluso controlar agregando datos e informaciones sobre las mismas. Así, por ejemplo, se olvida con bastante facilidad que los efectos de la escolarización en los individuos son siempre limitados y mediados por los efectos de otras organizaciones que se integran en estructuras sociales y culturales más complejas. Además, los procesos internos escolares se encuentran claramente mediatizados por las diferentes circunstancias e instancias culturales y sociales que rodean la vida de cada individuo. La escuela y el aula son sistemas complejos y dinámicos de individuos, tareas y relaciones, que no pueden entenderse mediante la simple agregación lineal de informaciones, se requiere elaborar un esquema o modelo complejo de comprensión. Lo mismo sucede con las interacciones entre el aula, la escuela y los contextos sociales y culturales en los que se ubica.

Impacto de las evaluaciones en los procesos de enseñanza

Como ponen de manifiesto múltiples investigaciones (Nevo, 2002), las evaluaciones externas relevantes, motivan a los docentes y a las administraciones a modificar sus prácticas de enseñanza y evaluación. Sin embargo, los cambios que motivan, y a veces imponen, suelen ser en general ajustes superficiales en los contenidos que hay que aprender y en las actividades que hay que desarrollar para preparar los test, y no cambios profundos en las prácticas de enseñanza y aprendizaje.

La información proporcionada por estas evaluaciones externas es fundamentalmente sumativa y útil en el mejor de los casos para identificar el progreso de cada escuela o del sistema pero de menor utilidad para orientar las prácticas de mejora, porque se distancian del currículo local y no proporcionan mucha información sobre las deficiencias y falsas concepciones de los estudiantes, sus fortalezas y debilidades y, sobre todo, porque nos dicen muy poco sobre sus causas.

Por otra parte, los ranking internos y externos, siempre adolecen de sobre simplificación de las diferencias culturales, sociales y nacionales. La publicación de ranking nacionales e internacionales puede fomentar la competitividad entre las escuelas en lugar de la cooperación, impidiendo uno de los requisitos de la innovación y mejora de las escuelas, la cooperación y aprendizaje mutuo (Goldstein, 2001). A este respecto conviene destacar que la utilización de los ranking de PISA puede reforzar las políticas educativas y la producción de estándares que contribuyen a normalizar, descontextualizar y empobrecer el curriculum escolar.

Con demasiada frecuencia, los test y pruebas de evaluación externa enfatizan los hechos y datos y no los conceptos, esquemas y modelos explicativos o interpretativos. Ahora bien, como ponen de manifiesto Biggs (2007) y Sunderman (2008) el aprendizaje y reproducción de los hechos no garantiza sino un aprendizaje superficial. El aprendizaje profundo y significativo requiere que los hechos se interpreten en base a modelos conceptuales, tomando en consideración el impacto de las variables sociales, históricas, políticas, lingüísticas y culturales que rodean la vida de los individuos.

Las pruebas y test de evaluación externa no pueden constituir por sí mismas herramientas de evaluación educativa. Miden o estiman resultados, indicadores de desarrollo, pero no diagnostican ni la naturaleza de los procesos, ni identifican las causas o factores que intervienen en la consecución de ese determinado grado de desarrollo, sus fortalezas y debilidades.

La evaluación educativa (Pérez Gómez, Soto Gómez, Sola Fernández y Serván Núñez, 2008) o formativa (Biggs, 2007) participa de la naturaleza constructivista del aprendizaje. En consecuencia, va más allá de los productos, incorporando los procesos y los contextos del aprendizaje. Pretende ayudar a reorientar los procesos cuando no funcionan adecuadamente, dar oportunidad a lo que se denomina *curriculum*¹ emergente, reformulando el diseño flexible de enfoque progresivo, para acomodarlos a las nuevas situaciones y a los actuales intereses y posibilidades de los aprendices.

La evaluación educativa², propone tareas y problemas auténticos en contextos reales y actuales, potencia la transparencia del sentido, al comunicar los propósitos y criterios de la evaluación; estimula y fomenta los procesos de reatualización, la evaluación entre pares y la propia autoevaluación, como la mejor estrategia para potenciar el desarrollo y la autorregulación del aprendizaje por parte del propio aprendiz. En definitiva, hace partícipe al estudiante de su propio proceso de aprendizaje.

Las evaluaciones externas basadas en amplias muestras y pruebas de papel y lápiz, por tanto, no pueden ni entenderse ni proponerse como evaluaciones formativas. Su pretensión legítima se circunscribe a proporcionar datos que ayudan a elaborar diagnósticos de situación, estimar las fortalezas y debilidades que demuestran los aprendices en ciertos ámbitos del saber, suponen llamadas de atención, identifican deficiencias, síntomas que hay que interpretar y en el mejor de los casos estimulan la reflexión de los implicados para indagar por debajo de tales síntomas y proponer estrategias de mejora. Constituyen una evaluación fundamentalmente sumativa que, en el mejor de los casos, puede inducir o desencadenar una evaluación formativa. Debe quedar bien claro que la evaluación formativa no puede confundirse con las evaluaciones externas basadas en pruebas estandarizadas, responde a otros propósitos, y requiere otros instrumentos, otras estrategias, y otros protagonistas³.

PISA, luces y sombras

Comprender las luces y sombras de este programa internacional de evaluación de los estudiantes requiere indagar la naturaleza y estructura de las pruebas, su diseño, desarrollo e interpretación desde los propósitos que el mismo programa establece como sus señas específicas de identidad

Los propósitos de PISA

PISA es un programa que desarrolla una comisión de la OCDE (Organización para Cooperación y Desarrollo Económico). Aunque parte de las experiencias previas de otros programas internacionales como TIMMS e IEA (International Evaluation Association), tanto PISA como IALS (International Adult Literacy Survey) no pretenden evaluar conocimientos vinculados estrictamente al *curriculum* escolar, pretenden identificar los aspectos comunes a los diferentes sistemas escolares. Se proponen, específicamente, estimar el grado de desarrollo en los niños y niñas de 15 años, de las aptitudes o competencias adecuadas para llevar a cabo las tareas necesarias en la vida cotidiana de la sociedad contemporánea. Se propone, por tanto como test independiente del *curriculum* nacional, regional o local. Un test intercultural e intercurricular que identifica las competencias necesarias para una vida satisfactoria en las complejas sociedades contemporáneas.

La fundamentación teórica de PISA se encuentra en el documento de la OCDE denominado DeSeCo (2002) (Definition and Selection of Competencies), donde se establecen y definen las competencias clave como las finalidades prioritarias de los sistemas educativos contemporáneos.

DeSeCo define el término de competencias como una combinación de conocimientos, habilidades, actitudes, valores y emociones que los seres humanos ponen en juego para afrontar los problemas complejos de la vida cotidiana que rodea su existencia. Son complejos sistemas de comprensión y actuación (Pérez Gómez, 2007, 2008, 2009).

DeSeCo propone como finalidades de los sistemas educativos contemporáneos tres competencias fundamentales:

- Competencia para *utilizar interactivamente y de forma eficaz las herramientas e instrumentos de todo tipo que requiere la sociedad de la información*, desde lenguajes hasta conocimientos (códigos, símbolos, textos, información, conocimiento, plataformas tecnológicas...) para comprender y situarse en el territorio natural, social, económico, político, cultural, artístico y personal. Los instrumentos y las mediaciones simbólicas componen la cultura contingente y relevante de la humanidad. Los construimos y nos construimos.

- Competencia para *funcionar en grupos sociales cada vez más complejos y heterogéneos*. El foco se sitúa en la interacción con “el otro”, con los otros diferentes e incluso discrepantes. Ello implica relacionarse bien con los demás, saber y querer comprender y cooperar, así como capacidad y voluntad para resolver con empatía y de forma pacífica y democrática los inevitables conflictos de la vida social.

- Competencia para *actuar de forma autónoma*. Lo que significa tanto el desarrollo de la propia identidad personal, como el ejercicio de autonomía relativa y con criterios propios a la hora de interpretar, decidir, elegir y actuar en cada contexto. Esta compleja competencia requiere: autoconcepto, autoestima y autorregulación, capacidad y voluntad para defender y afirmar los propios intereses y derechos, asumir las responsabilidades que se derivan de la libertad así como la capacidad y la voluntad para formar y desarrollar los propios proyectos de vida que incluye el ámbito personal, social y profesional.

Apoyándose en estas bases conceptuales iniciales, la OCDE sigue en la actualidad indagando la construcción compartida de las competencias fundamentales que requiere el ciudadano del siglo XXI⁴.

A partir de estos supuestos y principios, PISA pretende evaluar el grado de desarrollo de las competencias fundamentales que manifiestan los jóvenes de los diferentes países que participan en el programa.

Limitaciones internas

En sus documentos marco ya aparece el reconocimiento de la limitación que hace PISA del contenido de evaluación. Al utilizar por razones fundamentalmente de carácter logístico y de coste económico, un instrumento de papel y lápiz, PISA asume, al menos dos limitaciones sustanciales:

- 1.- Se limita a la evaluación de aquellos componentes de las competencias que pueden expresarse a través de dicho instrumento: conocimientos y capacidades. Por tanto en sus propios propósitos renuncia a evaluar competencias como sistemas complejos no solo de comprensión sino también de actuación, y por tanto quedan al margen los importantes aspectos emotivos, actitudinales y valorativos del desarrollo humano. A nuestro parecer, esto debe quedar bien claro, pues no se encuentra suficientemente explicitado en sus documentos y puede inducir a confusión. PISA no evalúa competencias, sino algunos componentes fundamentales de las mismas, en particular conocimientos y habilidades. Evaluar competencias requiere otros instrumentos y procedimientos de evaluación sensibles a la identificación y estimación de comportamientos, emociones y actitudes, para lo que se requeriría de otros procedimientos y herramientas de evaluación: observación de comportamientos y experiencias, portafolios, entrevistas, grupos de discusión y trabajo, proyectos...

Los ítems fácilmente se descontextualizan en una prueba de papel y lápiz, o se restringen artificialmente a contextos sólo relacionados con la vida cotidiana en alguno de los países implicados. Además, por exigencias psicométricas de claridad y pretendida objetividad, se excluyen aquellas preguntas de carácter controvertido y polémico desde el punto de vista científico, cultural, económico y político.

- 2.- Limita el objeto de evaluación a unos ámbitos determinados, que componen una parte de la primera competencia que define DeSeCo: La capacidad para utilizar interac-

tivamente los instrumentos simbólicos que ha ido creando la humanidad a lo largo de la historia y que se aloja principalmente en las disciplinas académicas. Dentro de esta primera competencia, PISA se ha circunscrito hasta el presente a cuatro ámbitos: Razonamiento matemático, comprensión lectora, pensamiento científico aplicado a las ciencias naturales y experimentales y solución de problemas. Queda fuera de la consideración de PISA un amplio territorio de las construcciones simbólicas del desarrollo disciplinar de la humanidad. Nada menos que el campo de las humanidades, las ciencias sociales, y las artes.

En el mismo sentido, cabe destacar que también quedan fuera de PISA el segundo y tercer ámbito de competencias fundamentales: -Capacidad de vivir y convivir en grupos cada vez más heterogéneos y la capacidad de actuación de forma autónoma, capacidad para aprender a aprender y a vivir. Por lo que debe quedar bien claro, cosa que tampoco aparece en los documentos marco, que PISA no mide el desarrollo de las capacidades y conocimientos que los jóvenes necesitan para incorporarse a la vida personal, social y profesional de las sociedades contemporáneas, sino solamente algunos aspectos importantes de dichas capacidades ya que al no considerar de partida estas áreas no sólo significa una restricción cuantitativa del objeto de evaluación, la desconsideración de una parte del territorio, sino que implica una importante restricción cualitativa, al prescindir de un componente ineludible de todos los problemas a los que se enfrentan los ciudadanos contemporáneos: la dimensión ética y de sentido como componente ineludible de las situaciones, las actuaciones y los propósitos de la vida humana individual y en grupo.

En los problemas y situaciones problemáticas que aparecen en las pruebas de PISA, se indaga siempre la dimensión técnica, la capacidad de razonamiento lógico y comprensión racional en los diferentes ámbitos, pero nada se alude a la dimensión ética de los mismos problemas o situaciones problemáticas que se abordan. Es cierto, como todo el mundo sabe, que afrontar la evaluación de la comprensión ética, acerca del sentido de las situaciones actuales o históricas implica afrontar cuestiones de valor, temáticas controvertidas y por tanto afrontar formas de evaluar y corregir más complejas y plurales.

Esta nómina tan abultada de ausencias aconseja una interpretación bien modesta y ajustada de los resultados como indicadores de calidad de los sistemas educativos evaluados.

En síntesis, los principios en los que se basa el programa de PISA: evaluar las capacidades de los estudiantes para entender y situarse ante los problemas complejos y cotidianos de la vida contemporánea y comprobar el grado de desarrollo de sus capacidades de razonamiento, análisis y comunicación, son fácilmente aceptables y asumibles por todos. No obstante la concreción de su objeto de evaluación, deja fuera importantísimas áreas del desarrollo cognitivo de los individuos y grupos humanos: Humanidades, Literatura, Arte, Ética, Filosofía, Ciencias Sociales, entre otras. Por ello debe declararse desde el principio, que PISA no mide ni estima la calidad de los sistemas educativos para el desarrollo integral de los individuos, sino aspectos muy concretos, aunque importantes, de los mismos.

3.- La tercera limitación autoimpuesta por los propios marcos conceptuales y propósitos de PISA hace referencia a la evaluación de conocimientos y capacidades con independencia del currículum escolar de cada país, región o localidad. Es una opción, a nuestro entender relevante y con sentido, pero que es necesario asumir con coherencia a lo largo de todo el proceso. Supone que con independencia de las concreciones culturales, sociales y políticas del currículum escolar, en cada ámbito objeto de evaluación es necesario ponernos de acuerdo sobre lo que entendemos que son los contenidos y capacidades esenciales para afrontar los problemas de la vida cotidiana de cada ciudadano del siglo XXI, situado a la vez en un mundo global y en un escenario vital, local, muy particular.

Los constructores de PISA parecen haber trabajado con dos supuestos básicos:

- El carácter universal de los contenidos y capacidades generales, con independencia de los contextos culturales y lingüísticos.
- El carácter estable y consistente para cada individuo de estas capacidades y contenidos que requiere su actuación eficaz en su propio contexto.

Esta pretensión de universalizar las habilidades o capacidades independientemente de los contextos donde se aplican puede ser contradictoria con la definición de competencias como sistemas complejos de comprensión y actuación en contextos concretos, así como del carácter situado de los conocimientos y habilidades, *situated cognition* (Brown, Collins y Duguid, 1989; Jonaert 2002; Lave y Wenger, 1991) que se propone como uno de los sustratos teóricos de la definición de competencias.

Partir de un marco teórico basado en competencias como sistemas complejos de comprensión y actuación, y, sin embargo, concretar las pruebas exclusivamente en algunos componentes de dichas competencias: capacidades y conocimientos, hace que aparezcan estas inevitables contradicciones. PISA argumenta la validez de sus pretensiones de universalizar tales habilidades y conocimientos, al incorporar en sus procedimientos de construcción de los test e ítems a expertos de diferentes contextos internacionales, y al exigir siempre la revisión de representantes de todos los países incorporados al programa, una de cuyas responsabilidades fundamentales es ajustar las adaptaciones de los problemas y situaciones a las peculiaridades de los contextos nacionales sin debilitar el carácter universal de las capacidades y conocimientos requeridos para su resolución.

En el rigor de estos procesos de revisión por parte del grupo de expertos internacionales reside gran parte de la credibilidad y validez de los instrumentos de PISA. La cuestión permanentemente pendiente es ¿mide PISA las capacidades que pretende medir, en cualquiera de los contextos en los que se está aplicando? ¿Algunas peculiaridades de cada contexto nacional, regional o local no pueden estar sesgando los resultados? ¿Si existe esa posibilidad de sesgos contextuales, son legítimas las comparaciones internacionales? Tales sesgos no invalidarían en absoluto las estimaciones locales y las posibilidades de compararlas consigo mismas a lo largo del tiempo, pero sí amenazan la posibilidad de comparación entre contextos y por tanto la elaboración de cualquier tipo de ranking.

Limitaciones externas

El comentario anterior nos lleva de lleno a una de las, a nuestro entender, limitaciones externas más serias de PISA y de todas las evaluaciones internacionales: La utilización política de los resultados en función del establecimiento de ranking y clasificaciones.

Hay pocas dudas en la actualidad sobre la enorme importancia política que ha adquirido PISA en la mayoría de los países implicados. En países como Noruega, Dinamarca o Alemania, parece haber provocado el efecto de un shock por lo inesperado de tales resultados, mientras que países como Finlandia, Corea, Singapur, Nueva Zelanda o Japón, se están considerando, por sus elevadas puntuaciones, como modelos de calidad en sus sistemas educativos que se proponen como ejemplos a estudiar e imitar (Carvalho, 2009).

Es evidente que el diagnóstico que hace PISA sobre el desarrollo de algunas capacidades y conocimientos de los ámbitos que estudia, no ofrece información útil suficiente para orientar políticas educativas que mejoren la calidad de los sistemas educativos. En todo caso, puede suponer un toque de atención sobre síntomas que será necesario indagar para comprender causas y factores explicativos. Por otra parte, la adopción de los países mejor situados en el diagnóstico de PISA como ejemplo a imitar, es una tentación simplista, por cuanto, los sistemas educativos y sus efectos, son subsistemas interdependientes, situados en complejos contextos culturales, económicos y políticos, que no

admiten comparaciones fáciles y requieren análisis complejos y detenidos de los múltiples factores diferenciales que provocan modos singulares de funcionamiento.

Las derivaciones políticas del establecimiento de ranking o de la propuesta de comparaciones fáciles e imitaciones simplistas de algunos aspectos aislados de los sistemas exitosos están condenados al fracaso y a la confusión al obviar el resto de variables complejas de carácter social, político y cultural que definen el contexto que enmarcan y posibilita cada uno de estos modelos.

Al mismo tiempo, la utilización política y mediática de los resultados de PISA, provoca interpretaciones simplistas y erróneas de la calidad de los sistemas educativos, que se concretan en imágenes extensamente difundidas por los medios y que acaban configurando la opinión pública de los ciudadanos, sin fundamento y sin matices. Por ejemplo, en España, los resultados deficientes en matemáticas y en comprensión lectora tanto en el 2003, como en el 2006, han provocado una, en parte, injustificada imagen de fracaso generalizado del sistema, y una utilización política interesada a favor de una forma más tradicional y academicista de exigir el esfuerzo de los escolares (*Back to Basic*), cuando lo que probablemente indican los resultados es más bien lo contrario: que una cultura excesivamente academicista del sistema educativo español y de los contenidos y procesos de enseñanza en torno a aprendizajes memorísticos de los libros de texto, no está preparando a nuestro alumnado para el tipo de razonamiento de problemas que plantean las pruebas de PISA en matemáticas y en lectura.

Es fácil comprender que en diferentes países o contextos culturales o políticos, los estudiantes reaccionan de distinta manera con distinta actitud, entusiasmo e implicación a las situaciones y problemas que presentan las pruebas de PISA y a la misma situación de la prueba. Por ello no es fácil establecer comparaciones ni ranking con cierto grado de legitimidad.

En segundo lugar hay que considerar como limitación externa, el origen y los intereses economicistas del programa de PISA. Los miembros de la OCDE se han implicado en PISA no por un interés general o científico en la investigación educativa o en las teorías del aprendizaje, la enseñanza o la evaluación. Su interés tiene un origen claramente económico que parte de un doble convencimiento:

- que la educación es un factor clave para el desarrollo económico y
- que la escuela actual no prepara para las exigencias cada vez más complejas del mercado laboral del siglo XXI, ni tampoco para las exigencias de la sociedad de la información.

Este origen de PISA en los intereses económicos de la OCDE puede impregnar y sesgar el tratamiento de los problemas educativos en general, así como la estructura, orientación y naturaleza de sus instrumentos y programas de evaluación. Por ello, se requiere una actitud vigilante para comprobar el deslizamiento de tales sesgos economicistas en el diseño, desarrollo o interpretación de las pruebas de evaluación, dentro del contexto de las exigencias de una economía de mercado global. PISA supone una opción, que responde en su origen a las exigencias de una economía liberal de mercado, y es necesario resaltar la existencia de opciones distintas desde el punto de vista cultural, político y económico, para diagnosticar los efectos de la escolarización actual en los estudiantes.

No obstante, esta necesaria prevención no puede utilizarse, *a priori*, como excusa genérica para rechazar o despreciar el valor ni de los instrumentos ni de los resultados. Se requiere un análisis concreto y matizado de las pruebas, los resultados y su utilización para comprobar a que propósitos y planteamientos específicos responden.

Fortalezas de PISA

En los intereses y propósitos de PISA y sobre todo en la claridad con que se expresan en sus marcos de referencia, residen gran parte de sus fortalezas:

PISA manifiesta reiteradamente su intención de medir o estimar el desarrollo en los jóvenes de 15 años, hacia el final de su educación obligatoria, de los conocimientos y habilidades que requiere su participación plena en los contextos complejos y cotidianos de la vida adulta. Este propósito supone por una parte distanciarse de la mera reproducción de los contenidos del currículum oficial, y por otra un esfuerzo por encontrar un fondo común, universal, de conocimientos y habilidades que independientemente de los contextos culturales, políticos y económicos, son requeridos en las sociedades interrelacionadas del mundo global en que se desarrolla la vida en el siglo XXI⁵.

PISA representa, por tanto, una apuesta decidida por la *utilización del conocimiento*, no por la mera reproducción de datos o conceptos acumulados en la memoria. En las pruebas de PISA la mera reproducción o repetición de datos o conceptos no se considera una capacidad digna de ser objeto de evaluación. Cuando para resolver problemas de la vida cotidiana como los que propone PISA, el sujeto necesita datos e informaciones, las pruebas y los ítems de PISA proporcionan dichos datos. Su preocupación, objeto de valoración, es la utilización de dichos datos en procesos y actividades de conocimiento de orden superior: aplicación, comprensión, transferencia, interpretación... En este aspecto PISA puede ponerse como ejemplo de evaluación que cambia y cuestiona drásticamente el modelo de evaluación habitual entre el profesorado de enseñanza secundaria y universitaria, al menos en España. Conviene no olvidar que la mera reproducción de datos y conceptos académicos acumulados en forma enciclopédica, y aprendidos de manera memorística, es la actividad predominante en la mayoría de las evaluaciones escolares en nuestro contexto español.

Como consecuencia del planteamiento precedente podemos destacar, en *segundo lugar*, que la independencia de PISA respecto a los currícula escolares, le permite situar al currículum y a las disciplinas que lo componen en el lugar que les corresponden. No pueden constituirse en las finalidades de los procesos de enseñanza-aprendizaje, como ahora ocurre, sino que han de ser considerados como medios y herramientas al servicio del desarrollo de las potencialidades humanas de cada individuo. El objetivo de la escuela no debe ser la enseñanza y aprendizaje del conocimiento disciplinar, sino el desarrollo de las competencias y cualidades humanas consideradas fundamentales.

La escuela convencional ha invertido de forma perversa la relación medios-fines: el aprendizaje de contenidos disciplinares y la superación de exámenes no pueden ser fines válidos en sí mismos, sino medios para facilitar el desarrollo de las cualidades o competencias humanas que consideramos valiosas. Si no lo consiguen, estos medios pierden su legitimidad (Pérez Gómez, 2009, p. 18)

PISA y la OCDE han hecho un esfuerzo encomiable por avanzar en ese camino, promoviendo grupos de expertos, foros y jornadas de trabajo para indagar y debatir el tema controvertido de las competencias clave que consideramos importantes para el ciudadano del siglo XXI (Carvalho, 2009).

En *tercer lugar*, tomando en consideración los propósitos de PISA y el modo de construir los ítems y pruebas, puede considerarse un modelo de evaluación que contempla algunas de las variables básicas de la evaluación auténtica, a pesar de que utiliza solamente instrumentos de papel y lápiz.

PISA... se centra más en la capacidad de los jóvenes de utilizar sus conocimientos y sus habilidades para hacer frente a los desafíos de la vida real que en saber hasta qué punto dominan un programa escolar concreto (OCDE, 2005, p. 20)

Los ítems de PISA pretenden medir o estimar el grado de utilización correcta que hacen los jóvenes de sus conocimientos y habilidades para afrontar problemas que se les presentan en la vida cotidiana. Este interés en situar los problemas en el contexto real de la vida cotidiana refuerza claramente la pretensión de PISA de acercarse en algunos aspectos a lo que hoy se denomina evaluación auténtica (Condemarin y Medina, 2000; Monereo, 2003, 2009; Wiggins, 1996).

Las preguntas que componen las pruebas pueden ser más o menos relevantes, más o menos abiertas y complejas, pero no cabe duda que en su mayor parte pueden considerarse significativas, aluden a situaciones que tienen sentido y significado para la mayoría de los estudiantes porque se refieren a problemas habituales y conocidos de la vida cotidiana.

Mención especial merecen las pruebas denominadas de resolución de problemas que se aplicaron en 2006, y que al parecer componen el sustrato de lo que se pretende que sea el horizonte de las pruebas de PISA en el futuro. En este caso, las preguntas, los ítems, aluden a problemas habituales de la vida cotidiana de mayor o menor complejidad, de carácter eminentemente interdisciplinar, para cuya resolución los jóvenes han de utilizar una importante y diversificada gama de conocimientos y habilidades, teóricas y prácticas. No obstante, conviene destacar en este sentido, que las pruebas de resolución de problemas versan sobre temas o situaciones de carácter eminentemente técnico, evitando, una vez más, los temas controvertidos, permeados de opciones de valor. Es decir, tanto en el contenido planteado, como en los criterios de valoración y corrección evitan las consideraciones de valor inevitables en los problemas sociales. Parece que se pretende descontaminar el conocimiento del inevitable compromiso social de su utilización real y concreta. Esta posición es especialmente llamativa en determinados estímulos, o tópicos-problema, donde temáticas como el sexismo, la desigualdad social, etcétera, no aparecen contemplados entre las respuestas válidas.

En *cuarto lugar*, cabe destacar la preocupación de PISA por identificar las *variables contextuales* que rodean la vida académica de los aprendices, e interpretar los resultados tomando en consideración la influencia mediadora de variables de carácter cultural, económico, social y de género. Además de aplicar cuestionarios sobre el contexto a alumnos y directores, PISA utiliza de forma clara las variables de contexto a la hora de interpretar los resultados y ofrecer interpretaciones más matizadas de los mismos. En consecuencia, establece *índices de equidad* de cada sistema educativo, a los que concede tanta importancia como a los índices que expresan el grado de excelencia de cada sistema. La calidad de la educación en cada país no puede estimarse ni única ni prioritariamente por el nivel medio de los resultados de la población estudiantil, es clave comprender el grado de dispersión de los mismos, lo que indica el grado de desigualdad-equidad del mismo. La educación obligatoria y gratuita en las sociedades democráticas se ha de preocupar fundamentalmente de extender los beneficios de la cultura a todos y cada uno de los individuos de cada generación. Además, los requerimientos del mundo del trabajo y profesional de la era contemporánea, a diferencia de la era pre e industrial de hace un siglo, suponen la extensión del conocimiento no sólo a las élites pensantes que se situaban en el vértice de la pirámide social, sino a todos los individuos que para producir, consumir y participar en la vida social requieren capacidades y conocimientos cada vez más complejos y elaborados.

En *quinto lugar*, puede destacarse como fortaleza el cuidado proceso de elaboración, aplicación, corrección y divulgación de los resultados de PISA. Constituyen ejemplos del rigor requerido en estos procesos para adquirir la legitimidad y la credibilidad necesaria para su aceptación, por parte de la sociedad en su conjunto, pero de forma muy especial por los directamente implicados, estudiantes y docentes. A diferencia de las pruebas de diagnóstico que están aplicando diferentes comunidades autónomas, las pruebas de PISA han utilizado un cuidado procedimiento para mantener la independencia, anonimato, control en la aplicación y homogeneidad en la corrección que garantizan la credibilidad requerida⁶.

En definitiva, podemos concluir que PISA constituye una valiosa herramienta de evaluación externa, sumativa, del desarrollo de algunos aspectos cognitivos de los jóvenes de 15 años, que no toma en consideración ámbitos fundamentales del desarrollo del ser humano, como el ámbito de las ciencias sociales, las humanidades y las artes, que no estima ni pretende estimar el desarrollo de competencias claves como la capacidad para

vivir en grupos cada vez más heterogéneos, o la capacidad de actuar de forma autónoma, que no mide ni estima competencias como sistemas complejos de comprensión y acción, sino capacidades cognitivas, que enfatiza capacidades de pensamiento no meramente reproductivas, que pretende trabajar con problemas auténticos en situaciones que representan condiciones reales y actuales y que supone un avance importante sobre los modos habituales y mayoritarios de evaluación del rendimiento académico de los estudiantes en el sistema educativo español.

Si bien es cierto que PISA ha adquirido un elevado nivel de impacto social, político y mediático, y que ha supuesto una importante llamada de atención sobre deficiencias reales de los sistemas educativos en la capacidad de lectura comprensiva, pensamiento matemático y conocimiento natural, pudiendo provocar y estimular movimientos de mejora, la utilización política interesada y simplista de índices y la cuestionable formulación de ranking y comparaciones internacionales sin considerar las complejas variaciones culturales, políticas y educativas de los distintos contextos nacionales o regionales, minan de forma significativa el valor educativo y político de los resultados de este omnipresente programa de evaluación.

Notas

- ¹ Utilizaremos el término *curriculum* en singular y *curricula* en plural respetando la forma latina, y por tanto sin tilde.
- ² El concepto de evaluación educativa que utilizamos en el presente artículo se encuentra estrechamente relacionado con los presupuestos que subyacen al movimiento de *assessment for learning* que puede consultarse en James *et al.* (2006), Black, McCormick, James y Pedder (2006), Black (2003), Hautamackky *et al.* (2002)
- ³ Como ejemplo de conciliación de las exigencias de evaluación externa con un proceso de evaluación interna, es interesante consultar el artículo de Simons en este mismo monográfico.
- ⁴ Puede consultarse al respecto las siguientes páginas web: <http://www.21stcenturyskills.org/> <http://www.pisaresconf09.org>
- ⁵ A este respecto es interesante consultar los desarrollos actuales del programa denominado las competencias del siglo XXI, en el que están implicados grupos de trabajo de la OCDE, estrechamente relacionados con los equipos de expertos que desarrollan PISA. Así como la página web con los contenidos de la conferencia internacional sobre los actuales desarrollos de PISA: <http://www.pisaresconf09.org>
- ⁶ Un ejemplo de las dificultades que a este respecto están encontrando las pruebas de evaluación de diagnóstico puede encontrarse en el artículo de Serván Núñez, en este mismo monográfico.

Referencias

- BLACK, P., HARRISON, C., LEE, C., MARSHALL, B. & WILLIAM, D. (2003). *Assessment for Learning - putting it into practice*. Buckingham: Open University Press.
- BLACK, P., MCCORMICK, R., JAMES, M. & PEDDER, D. (2006). Learning How to Learn and Assessment for Learning: a theoretical inquiry. *Research Papers in Education*, 21 (2), 119-132.
- BIGGS, J. (2007). *Teaching for Quality Learning at University*. Buckingham: Open University Press/McGraw Hill.
- BROWN, J. S. COLLINS, A. & DUGUID, P. (1989). Situated cognition and the culture of learning, *Education Researcher*, 18 (1), 32-42.
- CARVALHO, L. (2009). *Production of OECD's "Programme for International Student Assessment" (PISA) Knowledge and Policy*. Consultado en junio 2009: www.knowandpol.eu
- CONDEMARÍN, M. & MEDINA, A. (2000). *Evaluación auténtica de los aprendizajes. Un medio para mejorar las competencias en lenguaje y comunicación*. Santiago de Chile: Ed. Andrés Bello.
- GOLDSTEIN, H. (2001). Using pupil performance data for judging schools and teachers: Scope and limitation. *British Educational Research Journal*, 27, 433-442.
- HAMMERLEY, M. (2005). Is the evidence-based practice movement doing more good than harm? Reflections on Ian Chalmers' case for research-based policy making and practice. *Evidence and Policy*, 1, 85-100.
- HAUTAMÁKI, J., ARINEN, P., ERONEN, S., HAUTAMÁKI, A., KUPIAINEN, S., LINDBLOM, B., NIEMIVIRTA, M., PAKASLAHTI, L., RANTANEN, P. & SCHEININ, P. (2002). *Assessing learning-to-learn: a Framework* (Helsinki, National Board of Education) Helsinki: Helsinki University Printing House. Consultado mayo 2008, www.oph.fi/download/47716_learning.pdf
- JAMES, M., BLACK, P., CARMICHAEL, P., CONNER, C., DUDLEY, P., FOX, A., FROST, D., HONOUR, L., MACBEEATH, J., MCCORMICK, R., MARSHALL, B., PEDDER, D., PROCTER, R., SWAFFIELD, S. & WILLIAM, D., (2006). *Learning how to learn: tools for schools*. Londres: Routledge.
- JONNAERT, PH. (2002). *Compétences et socioconstructivisme. Un cadre théorique*. Bruselas: De Boeck & Larcier.
- LAVE, J. & WENGER, E. (1991). *Situated Learning: Legitimate Peripheral Participation*. Cambridge: Cambridge University Press.
- LINN, R. L. (2000). Assessment and accountability. *Educational Researcher*, 29 (2), 4-16.
- MONEREO, C. (2003). La evaluación del conocimiento estratégico a través de tareas auténticas. *Pensamiento Educativo*, 32, 71-89.
- MONEREO, C. (2009). *PISA como Excusa. Repensar la evaluación para cambiar la enseñanza*. Barcelona: GRAÓ.

- NEVO, D. (2002). Dialogue evaluation: Combining internal and external evaluation. En D. Nevo (Ed.), *School-based evaluation: An international perspective* (pp. 3-16). Oxford, UK: Elsevier.
- NORMAND, R. (2003). *Le mouvement de la school effectiveness et sa critique dans le monde anglo-saxon*. Bruselas: Université libre de Bruxelles.
- NORMAND, R. (2004). Les comparaisons internationales de résultats: Problèmes épistémologiques et questions de justice, *Éducation et Sociétés*, 12, 73-89.
- OCDE (2002). *Definition and Selection of Competencies (DeSeCo): Theoretical and Conceptual Foundations: Strategy Paper*. Consultado en http://www.statistik.admin.ch/stat_ch/ber15/desecco/desecco_strategy_paper_final.pdf
- OCDE (2005). *Informe PISA, 2003: Aprender para el mundo del mañana*. Madrid: Santillana. OCDE.
- ORIVEL, F. (2005). Foreword. En N. Postlethwaite, *Monitoring educational achievement* (pp. 10-14). París: United Nations Educational, Scientific and Cultural Organization.
- PÉREZ GÓMEZ, A. I. (2007). *La naturaleza de las competencias básicas y sus aplicaciones pedagógicas. Cuadernos de Educación 1*. Santander: Gobierno de Cantabria.
- PÉREZ GÓMEZ, A. I. (2008). ¿Competencias o pensamiento práctico? La construcción de los significados de representación y de acción. En J. Gimeno et al., *Educación por competencias, ¿qué hay de nuevo?* (pp. 59-103). Madrid: Morata.
- PÉREZ GÓMEZ, A. I. (2009). *Competencias, conocimiento y curriculum en la era de la información y de la incertidumbre*. Congreso Internacional de Investigación Educativa. 24 de Septiembre de 2009. Veracruz, México.
- PÉREZ GÓMEZ, A. I., SOTO GÓMEZ, E., SOLA FERNÁNDEZ, M. & SERVÁN NÚÑEZ, M^a J. (2008). *La evaluación como aprendizaje. Guías para el Espacio Europeo de Educación Superior nº 6*. Madrid: Akal.
- SUNDERMAN, G. (2008). *Holding NCLB accountable: Achieving accountability, equity & school reform*. Thousand Oaks, CA: Corwin Press.
- WIGGINS, G. P. (1996). Practicing what we preach in designing authentic assessments, *Educational Leadership*, 54 (4), 1-25.