

El precio de la evaluación estandarizada: la pérdida de calidad y la segmentación social

*The price of standardized evaluation:
loss of quality and social segmentation*

*O preço da avaliação padronizada:
perda de qualidade e segmentação social*

JUAN CASASSUS

Resumen: La calidad de la educación y sus concepciones y prácticas de evaluación son las preocupaciones centrales del autor del presente artículo. Su objetivo es mostrar como las pruebas estandarizadas referidas a norma, como las de los sistemas de evaluación centralizadas, no sirven para los propósitos que se le atribuyen socialmente, ya que no informan acerca de la calidad de la educación. El llama atención al poder desmesurado dado a los instrumentos estandarizados de medición en la evaluación y la gestión del sistema educativo, argumentando que, al contrario de lo esperado, ellos fomentan la desigualdad y disminuyen la calidad de la educación.

Palabras clave: evaluación y calidad de la educación; pruebas estandarizadas de medición; gestión educativa, segmentación social y pérdida de calidad.

Resumo: A qualidade da educação e suas concepções e práticas de avaliação são as preocupações centrais do autor deste artigo. Seu objetivo é mostrar como as provas padronizadas referidas a norma, como as dos sistemas centralizados de avaliação, não servem para os propósitos que se lhe atribuem socialmente, já que não informam sobre a qualidade da educação. Ele chama atenção ao poder desmesurado dado aos instrumentos padronizados de medição na avaliação e gestão do sistema educacional, argumentando que, ao contrário do esperado, eles promovem a desigualdade e diminuem a qualidade da educação.

Palavras-chave: avaliação e qualidade da educação; provas padronizadas de medição; gestão da educação, segmentação social e perda de qualidade.

Abstract: The quality of education and its conceptual formulations and practices on evaluation are the major concerns of the author of this paper. Its objective is to reveal how normative standardized tests, as those of centralized evaluation systems, do not attend their socially-intended purposes, since they do not inform on education quality. He calls attention to the excessive power given to standardized measurement instruments in the evaluation and management of educational systems, arguing that, contrary to expected, they promote inequality and hinder the quality of education.

Keywords: evaluation and the quality of education; standardized measurement testing; educational management, social segmentation, and loss of quality.

INTRODUCCIÓN

A inicios de la década de los 80, ocurrió un giro en la política de educación a nivel mundial que no tardó en llegar a América Latina. El giro consistió en que el foco de la política dejó de ser la expansión del sistema para concentrarse en lo que ocurría adentro del sistema. Se dijo entonces que estábamos pasando de un enfoque centrado en la cantidad hacia uno cuyo centro sería la calidad. Para ello, la nueva política se focalizó en cambiar la regulación del sistema, apoyada en la entonces naciente disciplina de la gestión educativa. Lamentablemente, las esperanzas que muchos tuvimos en ese entonces, se vieron frustradas, y tenemos que reconocer que las políticas que se concibieron en ese entonces, fracasaron en su intento. Un aspecto primordial de este fracaso, fue que nunca se dijo de lo que se entendía por calidad de la educación. De hecho nunca se debatió acerca de lo que era calidad y por lo tanto nunca hubo siquiera un intento de consenso. Es sorprendente que nunca se le dio un contenido a la palabra calidad. Igual que a principios del siglo XX cuando se debatió acerca de lo que era la inteligencia y la dificultad de definirla, se concluyó que inteligencia sería lo que midieran los tests de inteligencia, con calidad de educación se usó un símil: calidad se interpretó como siendo equivalente a un puntaje en una prueba estandarizada.

La ausencia de contenido y la ausencia de consenso acerca de la calidad, nos jugó una mala pasada. Por algún motivo aún no esclarecido se dio por sobreentendido que todo el mundo sabía de que se trataba la calidad de la educación, de tal manera, que cuando se formularon las políticas educativas desde esa época, bastaba con que se dijera que eran “para mejorar la calidad de la educación” para que fueran aceptadas por las autoridades. Sin embargo, éstas nunca fueron políticas de educación sino que fueron políticas de gestión. Las políticas educativas se centraron en la gestión del sistema, y nunca en la educación.

Solo para ilustrar el tema, recordemos algunas políticas destinadas a mejorar la calidad de la educación: mas horas de clases, calendarios mas extensos, descentralización, evaluación, (sí mucha evaluación estandarizada y centralizada), foco en los directores y su gestión, competencia entre escuelas, privatización, etc. Ninguna de estas políticas tan favorecidas por las autoridades nacionales e internacionales tuvieron que ver con educación. Tuvieron que ver con gestión del sistema. No se generaron en las disciplinas que se han ocupado de la educación y sus fines. No tienen su origen en la filosofía, ni la psicología, ni la pedagogía, ni en la sociología, sino en la economía.

Si miramos los criterios establecidos en estas políticas, podemos afirmar que las políticas puestas en practica desde hace 25 años han fracasado: los resultados que informan los sistemas de evaluación no solo no muestran progreso, sino en muchos países muestran retroceso (los puntajes bajan) y la segmentación social y la desigualdad social se ha fortalecido.

La política educativa es un arte difícil, porque la educación no es una ciencia precisa. Solo se la puede acercar con precaución y sentido común, pero también hay que estar dispuesto a cambiar. Ha llegado la hora de cambiar. Ha llegado la hora de usar el sentido común y hacer algunos planteamientos como los siguientes:

1. La finalidad de la educación (lo que yo entiendo como calidad de la educación) es que las personas sean mejores y con ello que la sociedad sea mejor. Esto implica reflexionar con profundidad acerca de las finalidades la educación en el Siglo XXI y cual es el tipo de persona, y cual es el tipo de sociedad que se quiere construir. Esto requiere terminar con el pensamiento simplista acerca de la educación que evita el debate de fondo y se queda en el manejo de herramientas.

2. Poner la educación al servicio de los niños, y no al servicio de la autoridad. Esto quiere decir que en vez de tener una educación cuya finalidad es el sometimiento de los alumnos al juicio moral de los mayores, se tienda a generar una educación que les brinde la experiencia del respeto y la convivencia armónica, que fortalezca su autonomía, que permita desarrollar su capacidad de reflexionar y pensar, que permita despertar el sentimiento de la empatía y la solidaridad, y sobre todo, que no les expropie su poder y dignidad personal.

3. Trabajar con los docentes y no contra ellos. Esto quiere decir que es necesario apoyar el trabajo de los docentes, escucharlos, confiar en ellos, solicitar de ellos las maneras de mejorar las cosas, darles tiempo y sobre todo no actuar por medio de presiones constantes cuyo efecto es deprimirlos y enfermarlos.

4. Considerar que son los procesos los que modulan las formas de la convivencia que las personas desarrollan consigo mismo y con los otros. Por ello es verdaderamente importante trabajar a nivel de los procesos que ocurren en las escuelas (interacciones y climas) y no pensar que el trabajo educativo reside en medir resultados.

El propósito de esta nota, es hacer un examen crítico de la política actual y elaborar las razones que nos llevan a pensar así. Estas razones se fundamentan en la evidencia empírica de que las políticas actuales han fracasado, que el énfasis en un tipo de evaluación deja al sistema atrapado en una dinámica perniciosa acerca del objetivo de mejorar la calidad y que tiene como consecuencia el efecto contrario al esperado, es decir que disminuye la calidad porque la empobrece en sus alcances y como instrumento de gestión expropia de los docentes la voluntad de mejorar la educación.

¿HAN FRACASADO LAS POLÍTICAS EDUCATIVAS PREVALECIENTES?

La política educativa de los noventa y las reformas educativas que proliferaron en América Latina (a la par con las de otras partes en el mundo) fueron un gran esfuerzo para movilizar las energías deprimidas por el costo que significó el ajuste estructural en el mundo de la educación durante la década de los ochenta. Desde los

inicios de los noventa, los gobiernos han aumentado los salarios docentes, construido nuevas escuelas, entregado muchos mas libros y se han fomentado nuevas leyes. Al mismo tiempo, se auto proclamó que el progreso en educación se reflejaría en el incremento de los puntajes de pruebas de opción múltiple.

A pesar de todos los esfuerzos desplegados mundialmente desde hace 20 años, los puntajes no suben en ningún país donde se mida el rendimiento académico. Las pruebas de los estudios comparativos internacionales (TIMSS, LLECE, PISA) así como las mediciones nacionales, que se basan en evaluaciones puntuales entregan resultados decepcionantes. De acuerdo a estos criterios, la educación parece estar estancada. Si aceptamos este principio, la evidencia empírica indica que los esfuerzos desplegados no han rendido los frutos esperados: en algunos países, como es el caso de Brasil, Chile, España, Francia o EEUU los puntajes han bajado y la desigualdad persiste. La respuesta que han dado las autoridades ante esta falencia no ha sido que es necesarios cambiar de política, sino mas bien, su postura ha sido la de profundizar en la política actual, situando en el centro de las políticas de estándares y mas evaluación cuantitativa. Tal opción, no puede sino empeorar la situación.

Mi planteamiento es que un factor crucial en el fracaso del sistema se sitúa en el sistema de gestión, y particularmente, en el sistema de evaluación y las teorías que lo subyacen. El sistema de evaluación oficial - que es el principal instrumento de gestión - se ha constituido en un freno para el desarrollo de la educación, genera involución y contribuye a disminuir la calidad de la educación. A continuación elaboramos sucintamente algunos argumentos que permiten fundar este juicio

EL USO DE LAS PRUEBAS ESTANDARIZADAS Y SUS LIMITACIONES COMO INSTRUMENTO DE EVALUACIÓN DE LA CALIDAD DE LA EDUCACIÓN

Las pruebas de opción múltiple, como las que utilizan los sistemas nacionales de evaluación como el SAEB en Brasil o el SIMCE en Chile son pruebas llamadas “pruebas referidas a norma”. El propósito de estas pruebas es determinar posiciones y establecer *rankings* entre los individuos que las responden. Este *ranking* se construye en relación a un promedio estandarizado, invariante y que es de 250 puntos. El creador de estas pruebas, Robert Glaser, ha sido muy claro en afirmar que estas pruebas no miden lo que saben los alumnos, ni lo que ellos saben hacer. Para evaluar lo que los alumnos saben o saben hacer se necesitan otro tipo de evaluaciones tales como los portafolios, pruebas de desempeño o pruebas de carácter constructivista. La única competencia que miden estas pruebas referidas a norma es, según Glaser (1963; 1997), la capacidad de recordar procedimientos (para las matemáticas) o reconocer un resultado cuando se le dan alternativas múltiples. Es un error conceptual decir que este tipo de prueba mide lo que saben y/o saben hacer los alumnos de las escuelas.

Es un error mas grave aún el equiparar el puntaje que se obtiene, con el objetivo de tener una educación de calidad. Puntaje no es calidad.

Calidad no es puntaje. Una educación de calidad es lo que puede llamarse una buena educación. Una educación de calidad tiene esencialmente que ver con la capacidad que tiene la institución escolar, de facilitar a que las personas se transformen en mejores personas, para que la sociedad se transforme en una mejor sociedad. Es una actividad de conocimiento transformativo. El nivel de calidad en la educación de una escuela es proporcional a la profundidad de los análisis que se ponga a la disposición de los alumnos, de los tipos de preguntas que se les permite proponer, del tipo de proyectos en los cuales se pueden embarcar y del tipo de problemas que son capaces de resolver. Calidad educativa no es una actividad centrada en tener puntajes mas altos. Son ámbitos distintos y es un error confundirlos porque tiene consecuencias negativas.

COMO EL PODER DESMESURADO DADO A LOS INSTRUMENTOS DE MEDICIÓN EN EVALUACIÓN AFECTAN NEGATIVAMENTE LA CALIDAD DE LA EDUCACIÓN

Voluntaria o involuntariamente se le ha dado a este tipo de evaluación la legitimidad y el poder de ser el eje articulador de la política educativa. Es a partir de los resultados de las pruebas de evaluación que se desprenden las políticas de *accountability*, de incentivos y castigos, de focalización recursos, de orientación de las capacitaciones, de uso de información para la toma de decisiones etc. Este es un contrasentido si lo contrastamos con las intenciones de mejorar la calidad de la educación. Si analizamos lo que ha ocurrido con la puesta en marcha de estos sistemas, veremos que trata de un mecanismo que en la práctica sirve para empeorar y no mejorar la calidad de la educación. De paso, vale la pena notar que es un error político el señalar que el éxito o fracaso de una política es subir los puntajes en una medición que no mide lo que se le atribuye, y que el sistema tal como esta concebido no puede modificar.

Uno podría preguntarse *¿porqué se afirma que el tipo de evaluación, instalada como medio de gestión, se ha convertido en un instrumento que baja la calidad, cuando esta diseñada para subir la calidad en educación?* La respuesta es que no solo se ha convertido en un freno al desarrollo de la educación, sino que contrariamente a lo que se afirma en los medios, esta involucionando el sistema. Esto se puede fundamentar desde distintos ángulos. A título de ejemplo podemos destacar que:

1. Los puntajes bajan. Como ya se ha señalado, la mera observación de las estadísticas comparables muestra que, a pesar de todos los esfuerzos, los puntajes siguen bajando. Ante esta situación, la tentación es profundizar una política de ejercer mas presión sobre los docentes, en vez de revisar el sistema.

2. Efectos perversos de las pruebas. La importancia que se le ha dado a esta forma de evaluación, hace que calidad de educación se haya transformado en

sinónimo de un puntaje en las pruebas. Esta es una identificación nefasta que tiene a su vez varios efectos:

- a) hace que el foco de la educación se ponga en la respuesta a pruebas psicométricas y no en la enseñanza, que es el fundamento de la profesión,
- b) saca de la vista cosas que son importantes como finalidades de la educación a saber: el desarrollo de la personalidad, el respeto, la ciudadanía, la curiosidad, la postura valórica, las ganas de descubrir conocimientos, compromiso con la sabiduría, etc. Todas cosas que no son evaluadas por las pruebas psicométricas. En definitiva, lo que es la finalidad de la educación queda de lado en el proceso educativo porque no es evaluado.

3. Subir puntajes crea una ilusión. identificar la calidad de la educación con subir los puntajes de las pruebas de medición genera situaciones ilusorias. Entrenar para responder una prueba estandarizada, es como cuando a una persona que tiene fiebre se la baja con medicamentos o enfriándole la cabeza, de manera que cuando se le pone el termómetro para medir la fiebre, ésta ha bajado, se concluye que el enfermo sanó. Igualmente, los puntajes se pueden subir o desvirtuar de muchas maneras, como por ejemplo:

- a) Inflando las notas (como lo muestran los sistemas de acceso a las universidades)
- b) Sacando a los “malos” alumnos cuando haya que dar las pruebas (cuando ese concepto de “malo” no tiene contenido específico sino que es sólo una interpretación)
- c) Entregando el resultado de las pruebas a los alumnos
- d) Entrenando a los alumnos a responder las pruebas.
- e) Los alumnos tienen la libertad de responder al azar

4. La educación se vuelve superficial. Focalizar la política en las pruebas instala una dinámica que hace que la educación se vuelva superficial.

- a) Porque la habilidad de responder a pruebas de alternativas múltiples no implica reflexionar de cómo alcanzar una respuesta, sino solo el acto elegir una respuesta. Aprender a reflexionar es importante porque es así como se logra profundizar un tema, y por ende elevar el nivel de calidad de dominio del contenido.
- b) Porque responder a pruebas de alternativas múltiples no requiere pensar ni razonar, sino reconocer una respuesta. Aprender a pensar y a razonar es importante porque permite encontrar las estructuras lógicas y a conectar los acontecimientos.

c) Porque responder a pruebas de este tipo, no requiere de la competencia de construir conocimientos, sino a la capacidad de “apuntarle” a una respuesta ya dada

d) “Enseñar para la prueba” implica que los profesores ocupan su tiempo en ejercitar a los alumnos a elegir. Implica que los profesores no le dedican tiempo a enseñar o a pensar.

e) Subir o bajar los puntajes, es una actividad superficial que no tiene que ver con una educación de calidad. Pensar que un sistema educativo mejora (o empeora) porque se suben (o bajan) los puntajes, es pesar que la salud es mejor cuando se le pone el termómetro a los enfermos después de haberles aplicado aspirinas y compresas de agua fría.

5. Consecuencias sociales: aumenta la desigualdad.

a) En los colegios de bajos recursos, enseñar para la prueba “para subir los puntajes ... cueste lo que cueste”, significa que los profesores se ven forzados a no educar, sino a entrenar a los alumnos para responder pruebas de opciones múltiples. En los sectores mas pudientes de la sociedad, esta actividad se paga fuera de las horas de clases.

b) Orienta el gasto hacia las pruebas en vez de la enseñanza. En la teoría subyacente a esta visión, se supone que el nivel de ingreso de la familia determina el resultado de la prueba. Las diferencias en los resultados de dichas pruebas no se deben a factores educativos. Entonces, como ya se sabe el resultado de antemano, y también siendo conscientes de que las variaciones son marginales, en vez de seguir evaluando (y pagando por ello) sería mas útil utilizar los recursos de la evaluación a actividades educaciones mas ricas y productivas para los alumnos.

EL ERROR DE USAR PRUEBAS ESTANDARIZADAS COMO INSTRUMENTO PARA LA GESTIÓN DEL SISTEMA EDUCATIVO

La evaluación en base a pruebas psicométricas referidas a norma está diseñada para seleccionar y no para incluir, por eso, en el contexto segmentado como los de los sistemas educativos en America Latinao el del sistema educativo chileno, fomenta la desigualdad. Ella no sirve para la mejorar la calidad de la educación, no sirve para que los profesores puedan hacer mejores clases, y habría que agregar que no son útiles para la gestión. Genera la ilusión de que se le esta tomando el pulso al sistema (ver la metáfora del termómetro mas arriba) pero solo habita en la irrealidad. Usarlas como instrumento de gestión solo sirve para identificar posiciones y rankings, o dicho de otra manera: segregar. Esto tiene como consecuencia generar mucho sufrimiento inútil e injusto en todos los niveles del sistema.

No hablamos de la evaluación en general pues es conveniente tener algún tipo de evaluación, sino de una evaluación que tiene como consecuencia, no mejorar la calidad sino aumentar la presión sobre los docentes mediante todo una parafernalia de premios y castigos. Es difícil comprender la lógica de las teorías subyacentes que llevan a pensar que los profesores van a enseñar mejor si se los castiga (o si se los premia). Es bien conocido en la literatura del “management” y en la práctica de la gestión de los ministerios y secretarías de educación, que este sistema no funciona. Algunas veces tiene un efecto de cortísimo plazo, pero este desaparece rápidamente.

Normalmente se piensa que los premios o los castigos son considerados incentivos para la acción y aparece como una medida fácil de implementar. Pero notemos que cuando se dice “incentivos” se está diciendo que se utiliza una motivación externa (extrínseca) por parte del sistema para que los docentes hagan su trabajo. Recién se mencionó que las motivaciones extrínsecas no funcionan para lo que quieren las autoridades del sistema. Sin embargo, sí funciona para otra cosa. La motivación extrínseca como instrumento y política de Estado tiene supuestos y efectos perversos.

1. Disminuye la dignidad del docente. El supuesto, es que supone que los docentes no están motivados intrínsecamente (internamente) para hacer su trabajo. Este es un supuesto falso y tiene un efecto perverso. Tenemos que preguntarnos: ¿por qué suponer que los docentes no habrían de estar motivados internamente a hacer bien su trabajo? Suponer que no lo están, como política de estado, es restarle dignidad al docente y a su labor. Esta pérdida de dignidad no se compensa con una mayor remuneración.

2. Mina la motivación intrínseca. Pero hay también otros efectos perversos, por ejemplo, suponer sistemáticamente que los docentes no están motivados para hacer su labor, mina sistemáticamente la motivación intrínseca. Una buena educación depende de la motivación intrínseca de los docentes. Tener una política que niega esta motivación intrínseca es una política que estimula una educación de mala calidad.

3. Rigidiza los procesos. Poner estándares y aplicarles evaluaciones, no solo fomenta la motivación extrínseca, sino que también rigidiza y estandariza los procedimientos. Por una parte, esto es algo que le conviene a los profesores que no están motivados para responder creativamente a los problemas que le presenta la enseñanza. Por otra parte, inhibe a los docentes que sí están motivados, y les “corta las alas” para encontrar las maneras más adecuadas de enseñar a los niños, en particular, cuando se trata de una política curricular constructivista.

4. Destruye el clima de aula. Enemista a profesores y alumnos y fragiliza la confianza y el vínculo entre ellos. La facultad de seleccionar alumnos hace que los alumnos “malos” sean expulsados hacia las escuelas municipales que están obligados a recibirlos. En un sistema convivial y sano, un alumno con dificultades es el desafío y la razón de la labor de los docentes en las escuelas más pobres. Pero en un sistema que opera sobre la base de juicios relacionados con los puntajes, donde la supervivencia de un docente depende de que los puntajes de sus alumnos, un alumno “malo” no puede sino ser visto por el docente como una amenaza y un problema que ojala “no existiese”.

CONCLUSIÓN

Para terminar esta nota, no podemos sino preguntarnos porqué hemos llegado a este estado de cosas. ¿Qué es lo que ha ocurrido que se abandonó la tradición de educar para la formación humana, de preocuparse por los alumnos, de educar para mejorar la sociedad? ¿Porqué se ha obligado a los docentes a abandonar el entusiasmo por su trabajo, haciéndolos entrar en el juego de las presiones y amenazas? Estas son preguntas que me he formulado y la dolorosa búsqueda de respuestas, me ha llevado a concluir que esto ha ocurrido porque la educación y muy en particular, la evaluación actual, se han transformado en un negocio muy lucrativo. Mientras sigamos pensando que el puntaje de una prueba psicométrica es una indicación de calidad, las autoridades estarán dispuestas a comprar sistemas de pruebas. Lo harán con entusiasmo con que se hace algo cuando uno cree que le están haciendo un favor a los mas desprotegidos. Creerán que cuando se los entrena en un sistema de instrucciones para responder las pruebas compradas a alto precio, se les estará ofreciendo una educación de calidad ... y no verán, como se dijo mas arriba, que se les estará entregando una educación empobrecida, pues el movimiento de los estándares y de pruebas precisamente se concentran en aquellos aspectos, como la memorización de información, que pueden ser medidos por dichas pruebas. Los estándares estandarizados y las pruebas psicométricas dejan de lado todo aquello que es mas difícil de medir: aprender a pensar, aprender a respetar, aprender a vivir con otros, aprender a hacer preguntas relevantes y resolverlas, a buscar la evidencia del conocimiento, a determinar qué es lo importante y valioso, a aprender del contexto. Es decir, todo aquello que parece estar en el corazón de una educación de calidad.

REFERENCIAS

GLASER, R. Instructional technology and the measurement of learning outcomes: some questions. *American Psychologist*, v. 18, 1963.

_____. Education for all: Access to learning and achieving usable knowledge. *International Seminar on Educational Evaluation*. Ministério da Educação, Rio de Janeiro, 30 de Novembro – 3 de dezembro de 1997.

■ JUAN CASASSUS é Diretor do Centro de Formação y Capacitación Índigo e do Mestrado em Educação Emocional da Universidad Academia de Humanismo Cristiano, em Santiago (Chile); Doutor em Sociologia da Educação pela Universidade de Paris V, Sorbonne; foi Especialista Principal de Educação da UNESCO para América Latina e o Caribe de 1992 a 2003. E-mail: jcasassus@vtr.net

*Recebido em março de 2007
Aprovado em abril de 2007*