

## La evaluación mediante pruebas de gran escala en México

### *Evaluation by large scale tests in Mexico*

**Edgar Jesús Bautista Sánchez**

Instituto de Ciencias de la Educación del Estado de México

[flobasatoluca@hotmail.com](mailto:flobasatoluca@hotmail.com)

### Resumen

El presente artículo es resultado del trabajo desarrollado al cursar los estudios de Doctorado en Ciencias de la Educación, siendo fundamental para integrar la investigación desarrollada sobre las tensiones que se generan en el docente de Educación Media Superior a partir de evaluaciones mediante pruebas de gran escala, que se han venido convirtiendo en el común denominador del sistema educativo de México en los últimos años.

El estudio desarrollado fue de tipo cualitativo y se realizó a partir de la obtención de narrativas a docentes, orientadores y directivos de un Centro de Bachillerato Tecnológico en el Estado de México, donde a partir del análisis desarrollado en el trabajo de campo se observa una tendencia por parte del docente a priorizar resultados de los exámenes de gran escala sobre los conocimientos del alumno, reflejándose en efectos colaterales de su práctica docente.

**Palabras clave:** Examen, evaluación de la educación, prueba de respuesta múltiple, maestro.

### Abstract

This article is the result of the work developed while studying the PhD in Educational Sciences, being essential to integrate the research carried out about the tensions that are generated in the teaching of Higher Education based on the evaluations using tests of large

scale, which have been the common denominator of the education system in Mexico in recent years.

The developed study was qualitative and was performed on the obtaining of narratives from teachers, counselors and principals from a Technical High School Centre in the State of Mexico, where from the analysis developed in the field, there is a tendency by the teacher to prioritize test results of large scale over the knowledge of the student, reflecting a collateral side effects on their teaching practice.

**Key words:** Examination, evaluation of education, test of multiple-choice, master.

**Fecha Recepción:** Octubre 2014 **Fecha Aceptación:** Diciembre 2014

---

## Introducción

El presente artículo se deriva del trabajo de investigación titulado “Tensiones generadas en los docentes de Educación Media Superior derivadas de la prueba ENLACE”, el cual fue desarrollado durante mis estudios de doctorado en el Instituto Superior de Ciencias de la Educación del Estado de México entre los años 2012 y 2014. Este artículo está conformado por tres apartados, el primero considera el origen y desarrollo de la evaluación, el segundo los elementos de la evaluación, el tercero refiere las áreas de tratamiento de la evaluación, el cuarto refiere el examen y las pruebas de gran escala nacionales e internacionales, y finalmente se presentan algunas consideraciones como conclusiones.

### Origen y desarrollo de la evaluación

Uno de los procesos que se relacionan de manera inmediata con el tema de la educación es la evaluación, que representa un término polisémico que se ha ido transformando y adaptando a las circunstancias a lo largo de los tiempos, propiciadas por las políticas públicas prevalecientes en ese momento.

El origen de la evaluación según Izquierdo (2008), ha sido objeto de un intenso debate sostenido por numerosos evaluadores como Guba y Lincoln (1989), quienes sitúan el nacimiento de la evaluación en la Norteamérica posterior a la Primera Guerra Mundial, vinculada a estudios de educación y salud pública, mientras que Patton (1990) la propone como área de investigación en las décadas de los años cincuenta y sesenta.

Las concepciones así como los modos de evaluar han ido cambiando conforme lo han hecho los diferentes paradigmas de la ciencia. Uno de los modelos iniciales de evaluación que predominó fue el modelo clásico, donde destacó Tyler y su evaluación por objetivos, para quien, como señala Izquierdo (2008), “el conocimiento de los objetivos es lo que va a guiar la elección de las variables dependientes y los criterios utilizados para la evaluación” (p.120). Carrión (2001), señala que Tyler lo denomina evaluación educativa, refiriéndola como el grado de cumplimiento de los objetivos que previamente se especificaban y se aplicaban. Sin embargo, ahora se observa que la evaluación actual ha dejado de fundamentarse en objetivos, buscando con ello el cumplimiento de diversas competencias.

En la década de los años cincuenta, Carrión (2001) señaló que todavía se consideraba que la evaluación en la educación se relacionaba exclusivamente con el aprendizaje del alumno, por lo que Benjamin Bloom propuso “complejizar” la idea de los objetivos educativos, al incorporar la psicología al desarrollo humano para hacerlos más integrales, lo que hizo surgir tres tipos de objetivos con correspondencia directa a tres dominios de la personalidad: cognitivo, afectivo y psicomotor. Cada uno de ellos está presente en los sistemas educativos de América Latina, en las competencias educativas pretendidas por sus sistemas educativos.

En los años sesenta, Carrión (2001) refiere que a partir de una gran cantidad de propuestas sobre los modelos de evaluación surge lo que se ha denominado investigación evaluativa, donde se propone que la evaluación sea un proceso sistemático de emisión de juicios, basados en una descripción de las discrepancias entre las ejecuciones de los estudiantes y profesores, sobre los estándares de aprendizaje.

Carrión (2001) refiere que en la década de los setenta Daniel L. Stufflebeam definió la evaluación educativa a partir de su relación con los procesos de planificación, como una

actividad que provee información para la toma de decisiones tales como: estructuración de procesos, ejecución y reciclaje, para dar origen a su modelo de evaluación que consideraba integrar el contexto, los insumos, los procesos y, finalmente, los productos. De entre estos, solo la evaluación del contexto era considerada como continua, mientras que las demás dependían de procesos de decisión. Aquí el contexto empieza a tomar un papel primordial dentro de la evaluación al considerar factores que la influyen de manera favorable o desfavorable.

Alicia de Alba (1991) señaló que en la década de los setenta “se empieza a hacer hincapié en la evaluación de programas, hasta considerarla como un área relevante dentro del campo. Es con el auge de la evaluación de programas que esta práctica se empieza a enraizar en los distintos sectores y ámbitos de la sociedad, de tal forma que comienzan a desarrollarse programas de evaluación institucional, sectorial y nacional” (p.80). Con esto dio inicio la evaluación de todo lo relacionado con la educación, involucrando tanto a alumnos, docentes, programas, currículos, instituciones, etcétera, todo ello bajo la pretensión de la tecnología educativa, que permitía validar y legitimar las decisiones que se tomaban sobre la evaluación, pero que escondía tras su discurso un problema de control, con implicaciones no solo en el ámbito educativo sino también en asuntos sociales en los cuales se empezaban a vislumbrar repercusiones en la disciplina y la percepción de los individuos cuando se enfrentaban a diversas problemáticas sociales.

Con respecto a la década de los años ochenta, Alicia de Alba (1991) refiere que Nilo hizo una consideración en la que señalaba que veía a la evaluación como una “herramienta delicada o muy afilada (axiológica)” (p.79), indicando que el principal problema específicamente para los países latinoamericanos radicaba en que no saben leer su realidad. Por su parte, él si podía percibirla porque veía dicha realidad desde fuera ya que pertenecía a uno de los países del grupo hegemónico-dominante de la OEA.

En esta década aparece nuevamente el problema de control. Se asume que los evaluadores en América Latina, los realizadores de operaciones, deben aplicar y presentar informes, cumpliendo funciones de técnico en evaluación. A partir de una serie de criterios y necesidades preestablecidos tienen como meta mostrar el estado de las cosas como algo natural. Alicia de Alba (1991) lo expresa puntualmente cuando señala “si la toma de

decisiones está en manos de algunos sectores o grupos, son estos los que ejercen el control; el evaluador toma como base una serie de criterios y necesidades ya construidos, y su trabajo se reduce a la operativización” (p.88).

Desde los años noventa en México:

Se promueven formas múltiples de evaluación en el ámbito educativo en términos de las instituciones, los académicos, los programas de investigación y docencia, y el aprendizaje escolar. Se aplica a todos los niveles educativos (desde preescolar hasta posgrado), y presenta diferentes manifestaciones tales como la acreditación, la certificación, el rendimiento de cuentas, el desempeño, los exámenes de ingreso y egreso... y desde diferentes planteamientos tales como la calidad, eficacia o la idoneidad. (Glazman, 2005, p.24)

Estos han incrementado los modelos referentes a la evaluación, tendencia de las políticas educativas globales, donde los países del grupo hegemónico-dominante señalan las estrategias a seguir, los parámetros a considerar, así como los métodos de evaluación y sus respectivas metas a corto, mediano y largo plazo.

Para el nuevo milenio, la evaluación prosigue con tratamientos aislados carentes de comunicación entre sus analistas, lo que ha ocasionado que la evaluación siga obedeciendo a demandas particulares en función de necesidades específicas. Ha mejorado las prácticas educativas, como señala Glazman (2005), “en términos de definiciones de los sujetos de evaluación, problemas suscitados, aclaración de criterios y en los propios informes” (p. 15), que respondan a esas demandas particulares y a necesidades específicas, dejando con ello relegada la formación de evaluadores especializados que pudiendo formar equipo con sus iguales para generar avances significativos, pero que en la parte de la evaluación se aíslan y siguen atomizando el conocimiento.

### **Elementos de la evaluación**

A partir de diferentes posturas de los evaluadores en las últimas décadas, el manejo de la evaluación ha tenido varios usos, llegándose incluso a reducir a algo técnico al priorizar la parte cuantitativa durante su proceso, situación que se ha venido reproduciendo en los

últimos años con influencia de la psicología conductual-experimental en campos técnico-administrativos de la sociedad, o en otras palabras, adscrita a un enfoque conductista y positivista.

Sin embargo, hay otra postura sobre la evaluación que es necesario recuperar, la del enfoque comprensivo que prioriza la parte cualitativa del citado proceso de la evaluación. Esta deja de centrarse en resultados y prioriza los procesos; reconoce desde un inicio que el evaluado es una persona y no un objeto, razón por la que se explicitará la evaluación a partir de la noción que maneja Alicia de Alba (1991). Ahí la noción de evaluación se constituye por dos elementos: la comprensión teórica-conceptual y la valoración axiológica.

#### a) Comprensión teórica-conceptual

Alicia de Alba (1991), señala que la evaluación en la educación es un “proceso complejo de reflexión, análisis crítico y síntesis conceptual valorativa, a partir del cual se conoce, comprende y valora” (p.94), este concepto parece distanciado de la manera como se ha estado manejando a la evaluación en las últimas décadas por el sistema educativo mexicano, en la que se minimizado o dejado al margen la interpretación humano-social, propia de la tradición aristotélica<sup>1</sup> (o del *verstehen-comprender*) y se ha dado una importancia cada vez mayor a lo funcional y mecanicista propio de las ciencias naturales y exactas de la tradición galileana<sup>2</sup> (o del *erklären*).

Considerar el elemento teórico de la evaluación es importante. Quien realiza la evaluación debe tener una comprensión teórica-conceptual de lo que se va a evaluar, es decir, realizar un análisis crítico desde una posición teórica que permita la comprensión de lo evaluado, ya que al carecer de ello caería en el enjuiciamiento y no la evaluación, situación que daría pie a toma de decisiones arbitrarias y autoritarias.

---

<sup>1</sup> Tradición aristotélica, exigía explicaciones que aclararan con qué fin ocurrían los fenómenos. Se decía que cada fenómeno tenía cuatro causas: formal, material, eficiente y final. Buscaba dar razón de los hechos al preguntar el ¿por qué? y el ¿para qué? de los mismos, priorizando el aspecto cualitativo.

<sup>2</sup> Tradición galileana. Se deja de mirar al universo como un conjunto de sustancias con propiedades para verlo como un flujo de acontecimientos que suceden según leyes, el centro deja de ser el mundo para ser el hombre. Prioriza lo cuantitativo, lo pragmático y lo útil.

Si considera el elemento teórico, la persona que realiza la evaluación dispone de los elementos mínimos necesarios para reconocer cada una de las partes que conforman al aspecto a evaluar. Por esa razón, si el evaluador evalúa sobre exámenes de gran escala debe tener una comprensión teórica-conceptual suficiente en lo que respecta a exámenes de gran escala para que le permita saber qué es lo que está evaluando.

b) Valoración axiológica.

El segundo elemento va dirigido a esclarecer las particularidades de la evaluación, y en donde intervienen valores de índole cultural, social, política, económica e ideológica, que finalmente conformarán el juicio con el que se realiza la evaluación, la cual es vinculada con la tradición aristotélica, donde la tarea de la evaluación es intentar dar respuestas a interrogantes como ¿para qué evaluar?, ¿quiénes evalúan?, ¿a quiénes se va a evaluar? Se refiere a un conjunto de valores a partir de los cuales se analiza el objeto-sujeto a evaluar y al compromiso que la instancia evaluadora adquiere con el sujeto, la condición o proceso que se encuentra en evaluación, pues finalmente el evaluador es un sujeto político que a través de esa tarea establece compromisos con quienes evalúan y con quienes se la han confiado.

El carácter axiológico de la evaluación se compone de distintas dimensiones que ayudan a integrar un juicio sobre lo evaluado, las cuales son: cultural, social, política, económica e ideológica, que desde la postura de Alicia de Alba (1991), se entenderían de la siguiente forma:

- Dimensión cultural: Se refiere a las diversas conformaciones culturales sobre las que se construyen las valoraciones, para evitar ajustar contradicciones culturales entre las instancias evaluadoras y las evaluadas.
- Dimensión social: Se refiere a la estructura social amplia (internacional, nacional, regional), así como a la organización social de los diversos grupos, sectores e instituciones.
- Dimensión política: Considera la valoración o rechazo por parte de la instancia evaluadora del proyecto político-social que habrá de desarrollarse, así como a la manera específica en la que afectan al proceso.

- Dimensión económica: Alude a la polémica entre estructura y superestructura, entre determinación, mediación y autonomía relativa, siendo esta una de las argumentaciones utilizadas para justificar los cambios en los sistemas educativos ya que en ellos se define el papel que jugarán los países subdesarrollados.
- Dimensión ideológica: Se da a la tarea de explicitar las diversas formas de concebir a la ideología como un conjunto de ideas acerca del mundo y que responden a intereses de cualquier tipo de conformación social que implique la cohesión de un grupo guiando su comportamiento hacia esos ideales.

Si bien todas estas cinco dimensiones del carácter axiológico de la evaluación manejan un conjunto de valores propios que tienen injerencia al momento de realizar la evaluación, cabe destacar que lo más importante que aportan a la misma es la relevancia que le dan a la contextualización como una parte estrechamente ligada al proceso de evaluación, pues evaluadores y evaluados son personas ante todo, vinculadas lo quieran o no a una realidad regida por estas dimensiones.

### **Áreas de tratamiento de la evaluación**

Existen dos grandes áreas de tratamiento de la evaluación, una pedagógica y otra administrativa, que a partir de los trabajos desarrollados por Raquel Glazman (2005), se perciben de la siguiente manera:

Carácter pedagógico.

De orígenes ancestrales, adquirió auge con el conductismo. En esta área, la evaluación tiene la finalidad de que el alumno aprenda y para ello se sustenta en elementos teóricos y conceptuales de la educación, siendo un proceso de esclarecimiento y búsqueda de datos que permiten precisar y ponderar situaciones o principios originados de un análisis del aprendizaje, para posteriormente corregir errores, revisar metas y cambiar en su caso propuestas metodológicas, con la idea de mejorar las condiciones de la educación.

Lo anterior se lleva a cabo si se cumple la siguiente serie de pasos considerados su soporte:

Fijar un objeto a evaluar.

Formular un juicio frente a un patrón.

Recoger la información.

Analizar la información.

Sistematizar la información.

Valorar la realidad evaluada.

Expresar dicho valor.

Y a partir de los pasos señalados se satisface la finalidad pretendida por el carácter pedagógico de la evaluación.

Carácter administrativo

Adquiere importancia en la década de los ochenta, donde la evaluación se asocia al rendimiento de cuentas, acreditación y certificación, y tiene como finalidad establecer jerarquizaciones, proponer comparaciones y tomar decisiones relacionadas con la inclusión y exclusión de sujetos, basadas estas en informes de las evaluaciones aplicadas, situación que se observará con detalle en las pruebas de gran escala.

Dentro de esta postura de trabajos relacionados a la problemática educativa destaca el enfoque socioeconómico de los sistemas educativos, funcionando bajo criterios económicos de producción, eficacia y control, que permitan preparar mano de obra calificada bajo la idea de elevar las tasas de desarrollo del país en cuestión.

Esta postura, según Glazman (2005), se relaciona con el carácter economista que ha asumido la política educativa de México y de los países en vías de desarrollo, donde se observa la reducción de los recursos gubernamentales destinados a la educación tomando como pretexto la crisis del sistema educativo correspondiente.

## **El examen y las pruebas de gran escala nacionales e internacionales**

Existen múltiples formas de evaluar y cada una de ellas necesita al menos de un instrumento para llevarla a cabo. En nuestro caso nos referiremos al instrumento de evaluación denominado examen y específicamente a los exámenes de gran escala.

Al pensar en un examen, las personas lo relacionan con acciones educativas y en específico con el proceso de evaluación, incluso se podría pensar que la misma relación ha formado parte de la historia del examen, sin embargo, Ángel Díaz Barriga (1993) demuestra la falsedad de dicha afirmación en tres puntos:

En primer lugar, el examen fue un instrumento creado por la burocracia china para elegir miembros de castas inferiores. En segundo lugar, hasta antes de la Edad Media no había evidencia de un sistema de exámenes ligados a la práctica educativa. En tercer lugar, la asignación de notas (calificaciones) es una herencia que el siglo XX dejó a la pedagogía, lo cual trajo consigo infinidad de problemas que todavía padecemos. (Díaz Barriga, 1993, p.13)

Díaz Barriga (1993) señala que “el examen es un espacio donde se realizan una multitud de inversiones de las relaciones sociales y las pedagógicas” (p.14). Específicamente hay tres inversiones: una que convierte los problemas sociales en pedagógicos, otra que convierte los problemas metodológicos en problemas de un solo examen y una última que reduce los problemas teóricos de la educación al ámbito técnico de la evaluación.

En términos de Foucault (2003), ello implica que el examen es un espacio que invierte las relaciones de saber y de poder. Las relaciones de saber en realidad son relaciones de poder. Esto ocasiona que la evaluación carezca de neutralidad y que entonces se propicie que los exámenes contemplen otra dificultad: que las calificaciones que genere difícilmente sean consideradas como objetivas y justas por su gran cantidad de subjetividades personales, académicas o institucionales. Lo anterior dificulta que este instrumento refleje todas las cualidades del aprendizaje, haciendo que esa calificación solo sirva para justificarse ante la sociedad y las autoridades asignadas.

Foucault (2003) refiere las repercusiones de poder y control inmersas en el proceso evaluativo, en su escrito “Vigilar y castigar”, donde las denomina como un conjunto de relaciones de sumisión, objetivación y normalización a partir de la escuela, al señalar que “el examen es la técnica por la cual el poder, en lugar de emitir los signos de su potencia, en lugar de imponer su marca a sus sometidos, mantiene a estos en un mecanismo de objetivación” (p. 174). Lo anterior se relaciona cuando Foucault (2003) considera que el examen combina dos tipos de técnicas, la primera es la de la jerarquía, utilizada para vigilar, y la segunda es la de la sanción, la cual normaliza, permitiendo a partir de ella calificar, clasificar y castigar. A pesar de ser este un estudio realizado por Foucault sobre exámenes de los siglos XVII y XVIII, puede extrapolarse y acoplarse a la realidad actual del sistema educativo mexicano.

Con respecto a las pruebas de gran escala, estas se desarrollaron en Estados Unidos a mediados del siglo XX, son pruebas tipo empleadas para evaluar los logros del aprendizaje en la parte conceptual y procedimental a un gran número de alumnos. Han sido recomendadas, para su uso en los sistemas educativos de un gran número de países, por parte de organismos internacionales, principalmente la Organización de Comercio y Desarrollo Económico (OCDE) y el Banco Mundial, donde la finalidad de dichos exámenes fue la de apoyar a los países en la mejora de su sistema escolar, buscando una mejor preparación de los jóvenes para su ingreso al mundo cambiante y globalizado. Como señala Maestro (2006), esto permite realizar comparaciones de los resultados para “valorar la eficacia y el funcionamiento de las políticas educativas adoptadas, el análisis de la evolución del sistema y su comparación con otras realidades educativas” (p. 315).

Entre las pruebas de gran escala de tipo internacional se encuentra la Program for International Student Assessment (PISA), en la que México, desde que ha participado, se ha mantenido en los últimos lugares. Tan solo en 2012, en matemáticas se ubicó en el puesto 53 de entre 65 países, mientras que en español se ubicó en el puesto 54 de entre 65 países.

Por lo anterior, si inicialmente se consideraba a los exámenes como un instrumento que apoyaba a la educación, hoy está iniciando una etapa en la que se priorizan los exámenes, en especial los exámenes de gran escala, implementando acciones que permitan elevar sus resultados a corto y mediano plazo, aunque ello no refleje realmente un incremento de los

conocimientos de los alumnos, pero sí de los indicadores que de estos se desprenden, pues esos resultados son ligados de cierta manera por autoridades educativas y medios de comunicación con el desempeño de los docentes e instituciones donde laboran.

La idea original del uso de las pruebas de gran escala se ha ido pervirtiendo como lo refiere Padilla (2009): “el examen con su aplicación masiva ha dejado de ser un instrumento del docente para convertirse en un instrumento contra el docente” (p.44). Se puede observar cómo de manera consciente e inconsciente se han estado implementando actividades como resolución de pruebas similares, buscando la familiarización a las mismas por parte de los alumnos, los repasos de temáticas abordadas en exámenes anteriores en tiempos de clase, etcétera. Con ello se da prioridad a resultados ligados con incentivos económicos y con la etiquetación del docente ante las autoridades, organizaciones públicas y privadas con injerencia en la educación. Por esta razón, en trabajos recientes realizados por Glazman (2005) se asume que las pruebas de gran escala reflejan y reproducen el posicionamiento económico y la ideología globalizadora de los grupos hegemónicos-dominantes, cuando consideran que la globalización “en su dimensión sociopolítica y económica adquiere connotaciones ideológicas específicas que determinan líneas de acción con efectos particulares en la educación” (p.19).

Dos de las críticas más recurrentes que se hacen a este tipo de pruebas son, por un lado, la descontextualización. Se asume que todos los estudiantes tienen los mismos propósitos escolares, mismas condiciones sociales y académicas, al utilizar exámenes iguales para alumnos con realidades diferentes. Por otro lado, está el ambiente de competitividad que generan estas pruebas entre alumnos, docentes, escuelas, estados e incluso entre países, a raíz de los rankings de las pruebas internacionales como PISA, o las aplicadas a nivel nacional en México como la Evaluación Nacional de Logros Académicos en Centros Escolares (ENLACE), que a partir del año 2014 suspendió su aplicación en el nivel de educación básica, principalmente por habersele ligado a prácticas no éticas durante el proceso de aplicación por parte de los profesores de dicho nivel educativo —se relacionó con estímulos económicos para los docentes de los alumnos con mejores resultados.

Se sigue aplicando en el Nivel Medio Superior bajo el argumento de que en este nivel no está viciado el examen desde la perspectiva de las autoridades de la Secretaría de Educación Pública.

Si bien en este nivel educativo no se relaciona la prueba ENLACE con incentivos económicos para el docente, existen puntos de alerta en el manejo de los resultados por parte de las autoridades educativas, que terminan por conflictuar al docente frente a grupo, ya que se prioriza el indicador de la prueba de gran escala frente a los conocimientos del curso.

### **Conclusiones a partir de los resultados de las pruebas de gran escala**

Como resultado de la implementación de las pruebas de gran escala dentro de los procesos de evaluación de los sistemas educativos y en especial el mexicano, se observan de manera recurrente los siguientes efectos:

1. Los resultados de las pruebas se han utilizado a conveniencia de organizaciones públicas y privadas.
2. Se priorizan los resultados de la prueba por encima del aprendizaje del alumno.
3. Se promueve la familiarización de estas pruebas de gran escala en los estudiantes.
4. Se ha convertido en un instrumento de control y rendición de cuentas.
5. El papel del docente se ha reducido a aplicarlos y entregar informes.
6. La estandarización en las pruebas asume alumnos iguales en realidades diferentes.

Los exámenes de gran escala en el ámbito educativo de México han sido desvirtuados de su intención original, al verse superados y priorizados por las exigencias de grupos hegemónicos dominantes representados principalmente por el Banco Mundial y la OCDE, propiciando paulatinamente sistemas educativos fallidos ante sus parámetros y criterios, provocando daños colaterales en los países que son miembros de estos Grupos u Organismos. En el afán de que estos países sigan teniendo privilegios primordialmente de tipo económico con inyecciones de capital a determinados sectores educativos, se seguirá priorizando al examen sobre el aprendizaje del alumno.

Las deficiencias señaladas de los exámenes de gran escala pocas veces se traducen en propuestas de mejora, debido a que simplemente se ha optado por implementar estrategias remediales que se traduzcan en periodos cortos de tiempo en la mejora de los indicadores, que finalmente solo reflejan una simulación o fantasía de un sistema educativo que es deficiente y que se desenvuelve únicamente para satisfacer los requerimientos del sistema hegemónico dominante.

## Bibliografía

Carrión, C. (2001) Valores y principios para evaluar la educación. México: Paidós.

De Alba, A. (1991) Evaluación Curricular. Conformación conceptual del campo. México: CESU-UNAM.

Díaz, Á. (1993) El examen. Textos para su historia y debate. México: Plaza y Valdés.

Ferrer, G. Arregui, P. (2003) Las pruebas internacionales de aprendizaje en América Latina y su impacto en la calidad de la educación: criterios para guiar futuras aplicaciones. PREAL. No. 26, pp.1-13.

Foucault, M. (2003) Vigilar y castigar. Nacimiento de la prisión. Argentina: Siglo veintiuno editores.

Glazman, R. (2005) Las caras de la evaluación educativa. México: PAIDEIA.

Guba, E. G., Lincoln (1989) Fourth Generation Evaluation. London: SAGE Publications.

Maestro, M. (2006) La evaluación del sistema educativo, en Revista de Educación, extraordinario. España: Ministerio de Educación y Ciencia, pp. 315-336.

Izquierdo, B. (2008) De la evaluación clásica a la evaluación pluralista. Criterios para clasificar los distintos tipos de evaluación, EMPIRIA revista de Metodología de las Ciencias Sociales, (16) pp. 115-134. Recuperado de <http://www.redalyc.org/articulo.oa?id=297124024005>

Padilla, R. (2009). Exámenes masivos internacionales y nacionales. ¿Encuentros o desencuentros?, *Perfiles Educativos*, 31(123), pp. 44-59. Recuperado 6 de noviembre 2013, de [http://www.scielo.org.mx/scielo.php?script=sci\\_arttext&pid=S0185-26982009000100004](http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S0185-26982009000100004)