

MODELOS TEÓRICOS E INDICADORES DE EVALUACIÓN EDUCATIVA

Por: Benilde García Cabrero

Currículo: profesora titular del Departamento de Psicología Educativa y Desarrollo de la División de Estudios de Posgrado de la Facultad de Psicología de la UNAM. Miembro del Sistema Nacional de Investigadores del Conacyt y consejera técnica del Instituto Nacional para la Evaluación de la Educación (INEE), donde colabora en la Dirección de Indicadores Educativos. Profesora de diversas universidades nacionales y evaluadora de programas en instituciones nacionales e internacionales, como UNESCO, CEREALC, UNICEF, OCDE, Conafe, IFE y SEP, entre otras.

Resumen

En este artículo se analizan los fundamentos teóricos en los que se sustenta el desarrollo de sistemas de monitoreo y evaluación de la calidad de la educación y se describen los modelos de evaluación más representativos del ámbito educativo. El artículo se propuso dar respuesta a las siguientes interrogantes: ¿qué implica monitorear y evaluar los sistemas educativos?, ¿qué papel desempeñan los modelos de evaluación, los indicadores y los estándares en estos procesos?, ¿cuáles son las fuentes de las que se derivan los indicadores educativos?, y ¿cómo se seleccionan los indicadores que pueden dar cuenta de la calidad de un sistema educativo?

Abstract

This article discusses the theoretical foundations that underpin the development of systems for monitoring and evaluating the quality of education and describes the most representative assessment models in education. The article pretends to answer the following questions: what does monitor and evaluate education systems imply? What role play assessment models, indicators and standards in these processes? What are the sources from which educational indicators are derived? How can the indicators that account for the quality of an educational system be selected? Key words: educational indicators, evaluation models, educational standards, monitoring systems.

INTRODUCCIÓN

El conocimiento y la información constituyen, de acuerdo con diversos especialistas, los elementos centrales de las diferencias entre las personas, las instituciones y los países en la época actual (Castells, 2002). Conforme a lo planteado por Mateo (2006), las sociedades postindustriales actuales se orientan, fundamentalmente, hacia la generación y gestión democrática del conocimiento. En estas sociedades, la informatización, la robotización y, en términos generales, la tecnologización, tanto de la industria como de la vida cotidiana, están propiciando el surgimiento de nuevas élites. El dominio de la tecnología para el manejo de la información y del conocimiento, así como su puesta en práctica, generan una distancia entre los grupos sociales que no tienen acceso a estos servicios y aquellos que los diseñan, utilizan o comercializan, con lo que producen situaciones de injusticia y marginación que atentan contra el bienestar y la cohesión social.

Ante este panorama, Mateo señala que la articulación de mecanismos socioeducativos que permitan que toda persona, en cualquier etapa de su vida,

pueda tener acceso a desarrollar instrumentos y competencias cognitivas que le ayuden a desempeñarse como un ciudadano activo y participativo que ejerce plenamente sus derechos, "...constituye un elemental ejercicio de justicia distributiva y de cohesión social y un objetivo político imprescindible" (2006, p. 243).

Frente a los enormes retos que plantea la sociedad postindustrial, la educación debe sistematizar sus esfuerzos y organizar sus recursos, de tal manera que pueda rendir cuentas en materia educativa sobre el cumplimiento de las metas que la sociedad le demanda. Como parte sustancial del proceso de rendición de cuentas, la educación debe instituir mecanismos que permitan evaluar sus recursos, procesos, agentes y resultados, y considerar, en la medida de lo posible, el conjunto del sistema educativo.

Para responder a las necesidades y demandas actuales en materia educativa, así como a los requerimientos de la globalización y la sociedad del conocimiento, el Programa Sectorial de Educación 2007-2012 propuso impulsar una profunda reforma educativa que se encuentra en marcha. Como parte de ella, se ha planteado promover la evaluación de todos los actores y procesos involucrados en el sistema educativo. Para ello, es conveniente supervisar el sistema; es decir, dar cuenta del progreso en el logro de las metas propuestas en el programa sectorial, así como del grado de aplicación de los mecanismos diseñados para alcanzarlas. Asimismo, definir y poner en operación sistemas de indicadores sobre la base del modelo explícito o implícito del funcionamiento del sistema educativo, para realizar evaluaciones diversas de componentes del sistema, o de éste en su conjunto.

MONITOREAR Y EVALUAR LOS SISTEMAS EDUCATIVOS

En el contexto de la definición de marcos de referencia para la evaluación de programas y proyectos, Dale (1998) establece una distinción entre dos conceptos: 1) monitoreo o seguimiento y 2) evaluación. El autor define el monitoreo como la recolección y el análisis de información "de modo rutinario y frecuente, acerca del desempeño o funcionamiento de un programa o proyecto. Esto puede hacerse a través de reuniones periódicas y presentación de informes o de investigaciones y estudios especiales. La información derivada de estos procedimientos debe ser vertida para realimentar el programa o proyecto, de preferencia para ajustar la etapa de planificación del ciclo y proponer acciones correctivas.

Scheerens, Glass y Thomas (2005) sugieren la utilización del término "monitoreo" en relación con la recopilación de información que se produce en el día a día, en el curso mismo de los acontecimientos educativos, y que sirve como base de las decisiones sobre la gestión. El monitoreo se apoya de manera fundamental en datos administrativos; en la realización de esta actividad existe una mayor preocupación por la descripción de la información recopilada que por la "valoración" propiamente tal (Scheerens, Glass y Thomas, 2005). El monitoreo de un sistema o programa educativo puede referirse tanto al del proceso como al de los resultados (Chen, 2005; Rossi, Lipsey y Freeman, 2004). Joo (2009) señala que el monitoreo del proceso es de gran ayuda para verificar que la implementación del programa sea conforme a lo planeado.

El monitoreo de los resultados se puede realizar a través de diferentes medios, en particular mediante la aplicación de exámenes que pueden funcionar como instrumentos para certificar estudiantes individuales, escuelas, zonas escolares, modalidades o tipos educativos, estados o países, y para regular lo que la sociedad

puede esperar de éstos, con lo cual se facilita la rendición de cuentas. El monitoreo de las características del sistema educativo puede realizarse considerando diversos niveles de agregación: el sistema educativo, un programa específico, una escuela, un salón de clases o una cohorte determinada de alumnos.

La evaluación, por su parte, se define como un examen más profundo, que se efectúa en determinados momentos de la operación de los programas o proyectos o de partes de ellos, por lo general con hincapié en el impacto, la eficiencia, eficacia, pertinencia, replicabilidad y sostenibilidad de dichos programas o proyectos (Dale, 1998). La literatura actual sobre el tema de la evaluación señala que las diversas aproximaciones desarrolladas hasta el presente coinciden en que su propósito es reunir información sistemática y hacer algún tipo de juicio sobre un cierto objeto de evaluación. En el ámbito educativo, estos objetos de evaluación pueden ser: los alumnos, las escuelas, el tipo de servicio educativo y el propio sistema educativo (Hansen, 2005; Stufflebeam, 2000). Scheerens, Glass y Thomas (2005) apuntan que una expectativa mayor respecto de los fines de la evaluación es que esta información ya “valorada” o “evaluada” pueda ser utilizada en condiciones ideales para tomar decisiones sobre el funcionamiento de los sistemas de educación, las escuelas, u otros agentes involucrados en el ámbito educativo; o de manera más amplia, en situaciones que impliquen la revisión o, incluso, el cambio del sistema que está siendo evaluado.

La evaluación del sistema educativo se apoya en la obtención de datos de diferentes fuentes; por ejemplo, los que se basan en medidas del rendimiento educativo; los que están disponibles en registros administrativos (incluyendo estadísticas educativas); y los que provienen de la revisión de expertos y la investigación educativa. De acuerdo con Scheerens, Glass y Thomas (2005), el propósito fundamental de la evaluación del sistema educativo es la determinación empírica de la calidad de éste. Las funciones principales de la evaluación son: 1) la certificación y acreditación; 2) la rendición de cuentas; y 3) el aprendizaje de la organización. La certificación y acreditación se dirigen, fundamentalmente, a precisar si las características del objeto evaluado se ajustan de modo formal a las normas y estándares establecidos.

La rendición de cuentas, por su parte, permite que la calidad del objeto pueda ser inspeccionada por otras instancias de la sociedad. Finalmente, el aprendizaje de la organización como estrategia de evaluación está dirigido a determinar si la evaluación de la calidad se utiliza como base de la mejora del objeto evaluado.

Estas tres formas de evaluación difieren en el grado de formalidad de los criterios y estándares que utilizan, en la naturaleza externa frente a la interna de los procedimientos de evaluación y en la orientación sumativa frente a la formativa que emplean. En el caso de la certificación y acreditación, la evaluación tiene un alto grado de formalidad; es sumativa y para desarrollarla se requieren estándares especificados para certificar estudiantes o profesionales y para la acreditación de programas. La rendición de cuentas requiere una evaluación formal de tipo externa con propósitos de control, en la que se combinan los enfoques formativo y sumativo y se emiten juicios en los que se prescriben acciones de sanción o recompensas en función de los resultados obtenidos. En el aprendizaje de la organización, la evaluación es menos formal y tiene carácter adaptativo; es más formativa que sumativa y se lleva a cabo con propósitos de mejoramiento mediante procedimientos de evaluación interna.

El desarrollo de planes sistemáticos de evaluación e indicadores nacionales

de la educación es una de las actividades establecidas en países como el Reino Unido, Holanda, Francia, Argentina, Chile, República Dominicana y México para la mejora de sus respectivos sistemas educativos. Asimismo, diversos organismos internacionales, entre los que destacan la OCDE, la UNESCO, la Unión Europea y la OIEA, han puesto en marcha programas y proyectos vinculados al desarrollo de las políticas de evaluación educativa.

La evolución y expansión de los sistemas de evaluación han implicado transformaciones en la concepción y práctica de la evaluación, particularmente cambios conceptuales en las nociones monolíticas de evaluación, y su sustitución por otras de carácter pluralista, así como el abandono de la idea de que la evaluación puede estar libre de valores. Se han introducido, también, cambios en las metodologías utilizadas, y ha surgido una creciente tendencia a la integración de métodos cuantitativos y cualitativos. Por su parte, *los usos de la evaluación se han modificado y se han destacado su carácter político y su capacidad como herramienta para seleccionar lo que se pretende evaluar y, por tanto, la posibilidad de influir en la orientación de los sistemas educativos.*

Se han introducido, de igual modo, cambios estructurales en la evaluación, que se caracterizan por su creciente inclusión en los mecanismos de gestión de los sistemas educativos, así como por una ampliación de los ámbitos de cobertura y una mayor interdisciplinariedad de la misma evaluación (House, 1993). Lo anterior pone de manifiesto la existencia de un nuevo fenómeno, que puede entenderse como la expansión del interés por la evaluación de los sistemas educativos, lo que ha tenido como efecto un rápido adelanto de la evaluación, concebida como disciplina científica y práctica profesional (Tiana, 2008).

ESTÁNDARES EDUCATIVOS

Como resultado del interés creciente por la evaluación, numerosos países, en particular los más avanzados, han adoptado estrategias del establecimiento de estándares para monitorear los resultados del sistema educativo a lo largo del tiempo (Tognolini y Stanley, 2007). Esta forma de monitoreo se basa en la determinación del crecimiento o avance de los estudiantes, en relación con resultados predeterminados para las diferentes asignaturas del currículo.

Los *estándares* pueden ser definidos como criterios claros y públicos que establecen los parámetros de lo que los alumnos pueden y deben *saber y saber hacer* en cada una de las asignaturas de los planes de estudio correspondientes a los diferentes niveles educativos. Los estándares constituyen una guía para que todas las instituciones escolares del sistema educativo cuenten con un referente sobre la calidad de la educación que se espera que ofrezcan a los alumnos y los resultados que deben alcanzar para lograr dicha calidad. Asimismo, los estándares sirven como marco de referencia para la actuación de profesores y directivos y para que los padres de familia y la sociedad puedan solicitar a las escuelas y al sistema educativo la rendición de cuentas sobre los resultados alcanzados.

Los estándares son afirmaciones explícitas del desempeño de los estudiantes que describen niveles de logro dentro de un área particular de aprendizaje. El documento "Standards, Assessment and Accountability" (Shepard, Hannaway y Baker, 2009) de la Academia Nacional de Educación de los Estados Unidos de Norteamérica señala que es necesario distinguir entre *estándares de contenido* y *estándares de*

desempeño. Mientras que los primeros se refieren al conocimiento y las habilidades que los estudiantes deben adquirir en una asignatura particular, los segundos constituyen ejemplos concretos y definiciones explícitas de lo que los estudiantes tienen que saber y ser capaces de hacer para demostrar su pericia en las habilidades y el conocimiento que están delineados en los estándares de contenido. Los estándares de desempeño se representan de mejor manera a través de muestras del trabajo de los estudiantes, que demuestran, por ejemplo, qué es lo que constituye calidad en un ensayo o plantean cómo se puede demostrar el dominio que se espera.

La mayoría de las veces, los estándares de ejecución se expresan simplemente como puntos de corte en una prueba (por ejemplo, 85% de respuestas correctas). Para determinar si los estudiantes logran los estándares, es necesario diseñar tareas de evaluación o exámenes acordes con dichos estándares. Éstos no están referidos a una norma y los reportes acerca del avance de los estudiantes no aluden a desempeños bajos, medios o altos, sino al porcentaje de estándares cubiertos por los estudiantes y al nivel de excelencia o calidad alcanzado, o bien, como plantea Strater (2006), a la profundidad y amplitud con que los lograron. Una ventaja de los sistemas de evaluación basados en estándares es su transparencia y la posibilidad de mejora continua que se deriva de su utilización.

Los términos “criterio de evaluación” y “estándar de evaluación” a menudo son confundidos (Scheerens, Glass y Thomas, 2005). El criterio es la dimensión en la que las interpretaciones evaluativas son realizadas; por ejemplo, un examen de matemáticas puede utilizarse como criterio en una evaluación educativa. El estándar se refiere, por una parte, al criterio (en el sentido que acabamos de definir) y, por otra, a una norma sobre cuya base puede decidirse si ha habido “éxito” o “fracaso”. Puntuaciones de corte definidas en una prueba de aprovechamiento particular son un ejemplo de estándares; en el caso del punto de corte, el estándar es absoluto.

En nuestro país se han establecido programas por competencias para los niveles de educación preescolar, primaria y secundaria y para algunas modalidades y sistemas de educación media y superior. Sin embargo, no se han establecido los criterios o estándares a alcanzar en relación con las competencias propuestas. En algunas asignaturas se han definido aprendizajes esperados, los cuales no alcanzan el estatus de estándar en vista de que no se han establecido los criterios de logro de dichos aprendizajes; por ejemplo, en la asignatura de Español de quinto grado se establece que los niños aprendan a redactar informes. Sin embargo, no basta con señalar que sean capaces de hacerlo, sino que se requiere mencionar, por ejemplo, el tipo de informes o las características mínimas que deben presentar.

INDICADORES EDUCATIVOS

De acuerdo con Scheerens, Glass y Thomas (2005), los indicadores educativos son estadísticos que permiten realizar juicios de valor sobre la pertinencia de los aspectos clave del funcionamiento de los sistemas educativos; constituyen características mensurables de éstos y aspiran a medir sus aspectos fundamentales. Proporcionan un panorama de las condiciones actuales del sistema educativo, sin describirlo a fondo y se espera que a través de ellos sea posible establecer inferencias acerca de la calidad de la enseñanza. Debido a lo anterior, los indicadores educativos tienen, en general, como punto de referencia un estándar contra el cual pueden efectuarse los juicios de valor correspondientes.

Los indicadores educativos deben ser susceptibles de comparación a través del tiempo, esto es, deben dar cuenta del progreso y los cambios en el desempeño de cierta variable; en otras palabras, el indicador no sólo cumple una función informativa, sino también evaluativa, pues se espera que, mediante ellos, sea posible determinar si existen mejoras o deterioros en algunas variables del sistema educativo (Morduchowicz, 2006). A este respecto, Kanaev y Tuijnman (2001), citados en Morduchowicz (2006), señalan que, además de su función informativa, los indicadores permiten construir nuevos enfoques y expectativas. Cabe destacar que los indicadores no sólo proporcionan información considerando el contexto del sistema, sino que facilitan el análisis de tendencias y la proyección de situaciones futuras del mismo sistema.

Según Morduchowicz (2006), desde la perspectiva de las políticas públicas, la selección y el uso de indicadores no es una tarea neutra; tiene dos dimensiones que permiten su análisis: la dimensión técnica y la dimensión política. Los indicadores pueden, entonces, referirse no sólo a las políticas, sino a las características generales del sistema educativo.

Los indicadores, en cuanto que constituyen propuestas de interpretación de la realidad, no pueden ser entendidos como herramientas capaces de proporcionar una visión acabada de la realidad tal cual es. En otras palabras, difícilmente permiten la comprensión de un fenómeno en toda su magnitud y complejidad, pues sólo aportan un marco de referencia cuantitativo que no incorpora los elementos cualitativos del fenómeno (Morduchowicz, 2006). Es comprensible, por tanto, que utilizar un solo indicador para obtener información acerca de un fenómeno sea inadecuado. Los indicadores no son cifras aisladas; se encuentran interrelacionados unos con otros, y para obtener una clara comprensión de la información que proporcionan deben agruparse y constituirse en lo que se conoce como sistema de indicadores.

Los sistemas de indicadores facilitan la descripción de situaciones que no pueden medirse de manera directa; por ejemplo, medir constructos tales como la calidad de la docencia o la de la educación puede resultar sumamente difícil, pues no existe un solo indicador que por sí mismo refleje este constructo. En cambio, si miden los mismos constructos considerando una serie de indicadores tales como formación académica, experiencia laboral, o resultados de aprendizaje y eficiencia terminal, sería posible configurar un panorama más claro en relación con estos constructos.

De acuerdo con Ogawa y Collom (1998), los sistemas de indicadores pueden adoptar dos características principales que hacen referencia al número de indicadores que conforman un sistema: indicadores parsimoniosos y extensos. Estos últimos incluyen un gran número de indicadores en su conjunto, y han sido criticados por resultar inmanejables y complejos; por su parte, los parsimoniosos contienen un reducido número de indicadores y las críticas que han recibido se relacionan, principalmente, con su poca capacidad para dar cuenta con eficacia de la complejidad de fenómenos que configuran al sistema educativo (Ogawa y Collom, 1998).

El sistema de indicadores educativos que publica la OCDE desde 1998 con el título de *Education at a Glance*, constituye un ejemplo de un sistema de monitoreo que pretende establecer comparaciones entre diferentes países respecto a la situación del sistema educativo. *Education at a Glance* incluye las visiones más actuales sobre la construcción y el cálculo de indicadores en diversos dominios educativos. Su publicación ha generado una amplia red de especialistas y un conjunto relevante de conocimientos (CERI, 1994), que ha impactado el desarrollo y

la evolución de los sistemas de evaluación y monitoreo en diversos países a escala mundial. En el contexto nacional se ha iniciado un proceso semejante a partir de la publicación anual del Instituto Nacional para la Evaluación de la Educación (INEE) denominada *Panorama educativo: indicadores del Sistema Educativo Nacional*, que se edita desde 2003 para presentar los resultados del diseño y la aplicación de un sistema de indicadores de contexto, insumo, proceso y producto.

UN MARCO DE REFERENCIA CONCEPTUAL PARA DESARROLLAR SISTEMAS DE MONITOREO Y EVALUACIÓN

Aun cuando existe un acuerdo generalizado entre los evaluadores de que el propósito de su práctica es determinar el valor de un objeto (Joint Committee on Standards for Educational Evaluation, 1994), existe un considerable desacuerdo respecto de qué significa valorar algo y cómo debe llevarse a cabo esta valoración. De aquí se desprende la multiplicidad de propósitos, perspectivas y metas, así como de los modelos utilizados para efectuar la evaluación (Donaldson y Scriven, 2003).

Este supuesto consenso respecto del propósito de la evaluación ha sido discutido por algunos evaluadores, que señalan que su responsabilidad fundamental es la descripción y explicación científica de las relaciones entre indicadores, y que el juicio de valor no es responsabilidad de los evaluadores, sino de los interesados. Esta perspectiva, denominada subjetivista o interpretativa, argumenta que los juicios de valor no son más que expresiones de preferencias personales o políticas, así como emociones o actitudes de individuos o grupos. Dichas expresiones deben distinguirse de aquellas basadas en los hechos que describen y explican alguna situación. Los hechos pueden ser debatidos y sus descripciones y explicaciones, juzgadas como verdaderas o falsas (Schwandt, 2005), pero no puede establecerse qué valor tienen. Por tanto, determinar la utilidad (efectividad-resultados) de un programa en función del logro de los objetivos que se propone, es el único "juicio" que puede ser "objetivo", porque la valoración descansa únicamente en los hechos.

Los defensores de la postura objetivista no están de acuerdo con la argumentación anterior y plantean que un juicio de valor (tal programa es bueno, pobre, etcétera) puede defenderse racionalmente. Dentro de la corriente objetivista existe un grupo que sostiene que sí es responsabilidad del evaluador emitir un juicio de valor, una vez que haya tomado en cuenta todos los aspectos que establecen el mérito, valía o significatividad del objeto evaluado, mediante una identificación de necesidades, la determinación del logro de estándares u objetivos del programa, o de comparaciones con programas alternativos, entre otros (Wholey, 2004).

Dentro de la corriente objetivista, pero desde una postura crítica, House (1993) y House y Howe (2001) señalan que la determinación del valor de un programa o política no puede ser realizada únicamente por los evaluadores, sino de forma conjunta por medio de algún tipo de procedimiento democrático o foro en el que los interesados y los evaluadores discutan y deliberen acerca del valor del objeto evaluado y lleguen a un consenso.

Las posturas contrapuestas respecto de si es o no responsabilidad del evaluador emitir juicios de valor como resultado de una evaluación, constituye uno de los temas álgidos de la discusión actual. La elección de un modelo de evaluación específico para sustentar las acciones de los evaluadores sobre un objeto de evaluación revela, en gran medida, la postura adoptada en esta discusión.

MODELOS TEÓRICOS DE EVALUACIÓN

Resulta de gran utilidad formalizar un proceso complejo como es la evaluación de los sistemas educativos, dentro de un modelo, que con frecuencia adopta la forma de un paradigma conceptual, un diagrama de flujo u otro tipo de representación esquemática (García-Cabrero, 1995, 1986). En la literatura sobre evaluación educativa existen diversos modelos de evaluación cuyas representaciones ayudan a examinar las relaciones entre diferentes componentes y procesos que ocurren dentro de dichos sistemas (Stake, 1967; Stufflebeam, 1968).

Madaus y Kellaghan (2000) mencionan que algunos autores han criticado la utilización del término modelo para referirse a las alternativas o perspectivas de evaluación. La crítica se fundamenta en el hecho de que estas perspectivas no cuentan con el nivel de formalización que tienen algunos modelos en ciencia, como es el caso de los modelos matemáticos. En respuesta y como defensa de la utilización del término modelo, los autores argumentan que las definiciones comunes de los diccionarios plantean que un modelo es una síntesis o abstracción de un fenómeno o proceso. Si se considera la evaluación como un proceso, entonces los modelos de evaluación constituyen la forma como un autor resume o conceptualiza la manera como debe efectuarse el proceso de evaluación. Un modelo puede definirse también como una metáfora, y en ese sentido, un modelo de evaluación puede concebirse como la metáfora o forma de pensar la evaluación que tiene el autor del modelo. Cada modelo se orienta a responder ciertas preguntas y, por tanto, para la elección de un modelo particular deben considerarse las preguntas que se pueden responder a partir de su utilización y los recursos que se tienen para responderlas. En situaciones de restricción presupuestaria es recomendable elegir varios modelos que permitan recolectar las mejores evidencias para responder las preguntas planteadas.

Madaus y Kellaghan (2000) sugieren también la posibilidad de combinar los datos obtenidos mediante: 1) la aplicación de pruebas de rendimiento que se utilizan en el modelo de evaluación orientado a las metas; 2) los recursos asignados, considerados dentro del modelo orientado a las decisiones; y los 3) datos provenientes de observaciones y entrevistas que típicamente se aplican en los modelos naturalistas. Los autores afirman que la combinación de modelos constituye la mejor vía para documentar la complejidad de los sistemas y programas educativos.

La gran cantidad de modelos en el ámbito de la evaluación educativa revela el amplio rango de posiciones epistemológicas e ideológicas que existen entre los teóricos sobre la naturaleza de la evaluación, la forma de conducirla y la manera de presentar y utilizar los resultados.

DE LOS MODELOS TEÓRICOS A LA CONSTRUCCIÓN DE INDICADORES DE EVALUACIÓN

Diversos autores coinciden en que, para desarrollar indicadores, se debe adoptar un modelo del sistema educativo (Blank y Gruebel, 1993; Jones y Nielsen, 1994). Los modelos permiten identificar los elementos clave que requieren ser medidos dentro del sistema; incluyen las relaciones teóricas entre los componentes y permiten distinguir áreas en las que es necesario intervenir. Los sistemas de indicadores en su conjunto se derivan de los modelos de evaluación. La operacionalización de los componentes del modelo en entidades mensurables es lo que se conoce como indicadores; en tanto que el conjunto de indicadores constituye un sistema (Ogawa y Collom, 1998).

Hansen (2005) efectuó una revisión de dos enfoques alternativos de evaluación. El primero consiste en el diseño y conducción de un estudio para proporcionar evidencia acerca del valor y mérito de un objeto particular de interés para una audiencia específica. El segundo, en una valoración retrospectiva y cuidadosa del mérito y valor de la administración, productos y resultados de intervenciones gubernamentales dirigidas a desempeñar un papel en situaciones prácticas de intervención (Vedung, 1997). La primera aproximación es conocida como evaluación de programas (Stufflebeam, 2000; Alkin, 2004) y la segunda, como evaluación de la efectividad organizacional (Scott, 2003).

Tanto Payne (1994) como Hansen (2005) mencionan que los modelos básicos de evaluación utilizados dentro de los dos enfoques se traslapan de forma significativa. Aunque no existe un consenso en la literatura sobre la clasificación de los tipos y modelos de evaluación que existen, es posible agruparlos en seis grandes categorías: 1) modelos de resultados; 2) modelos explicativos del proceso; 3) modelos económicos; 4) modelos de actor; 5) modelos de teoría del programa; y 6) modelos sistémicos (Hansen, 2005).

El modelo de resultados, también denominado modelo basado en las metas, se centra en los productos alcanzados por un programa u organización; tiene dos subcategorías: el modelo de logro de metas, y el modelo de efectos. El primero es un clásico en la literatura sobre evaluación de programas y evaluación organizacional; en él los resultados se evalúan con base en las metas u objetivos establecidos, tal como lo señalan: Popham, 1970 (citado en Escudero, 2003); Scriven, 1973 (citado en Escudero, 2003); Stainmetz, 1983; Tyler, 1969 (citado en Escudero, 2003). Por su parte, el modelo de efectos tiene la intención primordial de conocer todas las consecuencias producidas por el objeto que se evalúa. Este modelo, también conocido como “modelo de evaluación libre de metas” (Scriven, 1973, citado en Escudero, 2003), ha sido criticado porque se considera que puede tener criterios de evaluación deficientes.

El modelo explicativo del proceso se centra, como su nombre lo indica, en los procesos y esfuerzos. La evaluación de proceso se lleva a cabo de manera adecuada cuando se realiza “en tiempo real”, y es menos pertinente cuando ocurre a través de análisis históricos. Los modelos económicos consideran al objeto de la evaluación —el programa o la organización— como una caja negra que relaciona la evaluación de los resultados (ya sea en forma de rendimiento de la producción, efectos o beneficios más duraderos) con los insumos (entradas) (Hansen, 2005).

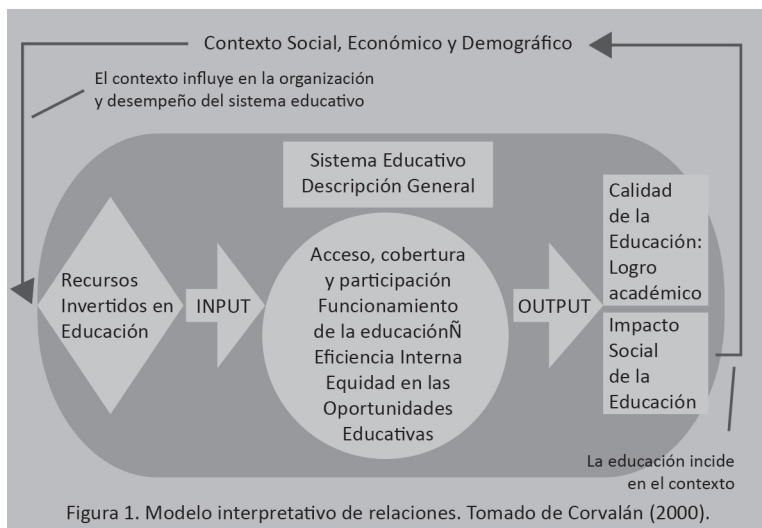
Los modelos de actor (clientes, interesados y pares) se basan en los criterios establecidos por los propios actores de la evaluación. El primero se centra en los criterios de los clientes para llevar a cabo la evaluación; el modelo de los interesados considera los criterios de todas las partes interesadas que resultan pertinentes para la evaluación (Stake, 1967, 2006). Por su parte, el modelo de revisión por pares toma en cuenta los criterios para la certificación de profesionales. El modelo de teoría del programa se centra en la evaluación de la validez de la teoría en la que se fundamenta el programa, una intervención u organización determinados. Este modelo compara y reconstruye, con base en el análisis empírico (Birckmayer y Weiss, 2000), las relaciones causales entre el contexto, el mecanismo y los resultados (Pawson y Tilley, 1997; Pawson, 2002). El modelo de teoría del programa puede ser visto como un modelo de resultados extendido. El objetivo del modelo teórico es revisar y seguir desarrollando la teoría del programa y así aprender lo que funciona, para quiénes y en qué contextos.

Por último, el modelo sistémico aborda el análisis de los insumos, la estructura, el proceso y las salidas en términos de resultados (Sheerens y Creemers, 1989). En él, la evaluación se fundamenta tanto en comparaciones de insumos planeadas y realizadas como en los procesos, estructuras y contrastación de resultados de programas similares u organizaciones reconocidas como de excelencia. Este modelo ha sido ampliamente utilizado en la práctica de la evaluación, y como ejemplo puede citarse el modelo CIPP (contexto, insumo, proceso, producto) desarrollado por Daniel Stufflebeam (Stufflebeam y Shinkfield, 1987; Stufflebeam, 2004), en el que se basa el modelo del sistema de indicadores del Instituto Nacional para la Evaluación de la Educación (INEE, 2003, 2004, 2005, 2006, 2007 y 2008).

En las últimas cuatro décadas, en particular en Estados Unidos de Norteamérica, se ha desarrollado un número importante de modelos de evaluación; algunos de ellos, como el modelo de orientación a metas y los modelos económicos, se apoyan en la noción de objetividad y consideran que la evaluación puede producir resultados objetivos. En algunos otros modelos de corte cualitativo/interpretativo, como el de los actores, los criterios de evaluación no se elaboran de forma previa, sino que se formulan durante el propio proceso de evaluación; en éstos no se considera que la evaluación produzca resultados objetivos, ya que cada actor o participante tiene perspectivas o percepciones diferentes del objeto evaluado.

Todos los modelos tienen tanto fortalezas como debilidades, y la elección de alguno de ellos corresponde, de acuerdo con Hansen (2005), a la decisión de elegir una perspectiva del objeto de estudio. Al seleccionar un modelo, algunos aspectos del objeto evaluado pueden ser enfocados claramente, mientras que otros se excluyen del foco.

A continuación se presenta un ejemplo de cómo se puede utilizar el modelo sistémico para derivar indicadores a partir del modelo interpretativo de relaciones de Corvalán (2000) (ver figura 1), el cual considera el contexto en el que se desenvuelve la educación y la forma en que ésta se organiza. Incorpora los insumos destinados a la educación medidos en términos de recursos humanos, materiales y financieros y centra la atención en el funcionamiento de la educación propiamente tal. En este modelo, los resultados de la educación son concebidos en cuanto al logro académico e impacto social, que, a su vez, influyen de nuevo en el contexto.



A partir de este modelo de relaciones se pueden derivar cinco grandes categorías de indicadores: 1) contexto demográfico, social y económico y descripción general del sistema de educación; 2) recursos en educación; 3) desempeño de los sistemas educativos; 4) calidad de la educación; y 5) impacto social de la educación.

Contexto demográfico, social y económico y descripción general del sistema de educación

Estos indicadores, según Scheerens, Glass, y Thomas (2005), son en general definidos en el nivel de los sistemas educativos nacionales, y se refieren básicamente a las características de dimensión y estructurales del sistema educativo nacional. Algunos ejemplos pueden ser: metas educativas y estándares por nivel educativo; edad relativa de la población estudiantil; la estructura de las escuelas en el país, entre otros.

Recursos en educación

Permiten conocer la disponibilidad, características generales, técnicas y distribución de los recursos humanos, materiales, así como financieros que contribuyen al proceso educativo. Algunos ejemplos de este tipo de indicadores son: gasto educativo por estudiante; gastos en desarrollo e investigación educativa; o porcentaje de la fuerza laboral del país que se dedica al campo educativo.

Desempeño de los sistemas educativos

Dan cuenta de las características del entorno de aprendizaje, así como de los aspectos de organización de la escuela. Algunos ejemplos son: los patrones de centralización/descentralización del sistema; el tiempo de enseñanza por asignatura; la proporción del presupuesto total dedicado a la educación en relación con reformas a programas específicos; la inversión y los arreglos estructurales para el monitoreo y la evaluación del sistema educativo, entre otros.

Calidad de la educación

La preocupación por asegurar la calidad y equidad de la educación en una concepción amplia y diversa, es uno de los principales factores que determinan el desarrollo de este tipo de indicadores. Se relacionan principalmente con estadísticas sobre el acceso y participación, y sobre todo con los datos que refieren el logro educativo. Algunos ejemplos son: tasas de participación en los diversos niveles educativos; progresión a través del sistema educativo; promedio del logro en dominios curriculares básicos; competencias transversales; y habilidades para la vida (habilidades sociales, resolución de problemas, entre otras).

Impacto social de la educación

Permite el análisis de la relación entre la educación y la sociedad. Este tipo de indicadores se refiere a los cambios que ocurren en los diversos sectores sociales y que pueden ser interpretados como efectos de la educación; por ejemplo: el impacto de la educación en el empleo/desempleo de la juventud; el posicionamiento de los egresados con cierto nivel de certificación en el mercado laboral; ingresos vinculados a la educación y el nivel de entrenamiento de los individuos, entre otros. Un ejemplo del uso de este tipo de indicadores se presenta en el análisis de Muñoz-Izquierdo y Márquez (2000) respecto del impacto que tuvo la expansión de oportunidades educativas durante las últimas décadas del siglo anterior, sobre las ocupaciones desempeñadas por egresados de los sistemas educativos.

Un aspecto que debe considerarse para desarrollar sistemas de indicadores es el número de éstos que deberá incluirse. A este respecto, la literatura señala que los sistemas extensos de indicadores pueden resultar inmanejables y demasiado complejos, con lo cual se volverían poco útiles (Blank, 1993; McDonnell y Oakes, 1989; Shavelson, McDonnell y Okes, 1989). Por consiguiente, es deseable que prevean sólo un número esencial de indicadores. No obstante, aún existe una falta de consenso acerca de la cantidad de indicadores que puede considerarse aceptable y funcional dentro de un sistema de indicadores en particular. Algunos autores, como Dickson y Lim (1991), plantean que se deben incluir entre cinco y ocho; otros (por ejemplo, Blank, 1993), que basta con doce, mientras que algunos como Hafner y Buchanan (1992) recomiendan un máximo de veinte.

Resulta evidente que el número de indicadores debe tomar en cuenta los rangos que se establecen en la literatura, sin que esto se convierta necesariamente en una regla de oro que no debe ser quebrantada. Al tomar la decisión del número de indicadores a incluir, se debe también considerar el contexto particular en el que se inscribe el sistema, y en función de éste, precisar el número de indicadores que lo constituirán.

El sistema de indicadores desarrollado por la OCDE en 1998 incluyó 38 indicadores, agrupados en seis categorías: 1) contexto demográfico, social y económico en el que operan los sistemas educativos; 2) recursos humanos y financieros invertidos en educación; 3) acceso, participación, progresión y eficiencia terminal en educación; 4) participación de los jóvenes en educación (quince a veintinueve años), así como seguimiento de la eficiencia terminal en educación inicial; 5) entorno de aprendizaje y formas en las que los sistemas educativos se organizan para mostrar datos en relación con los profesores y la enseñanza; y 6) resultados individuales y sociales del mercado de trabajo de la educación (OCDE, 1998). En 2009, los indicadores de la OCDE se redujeron a 25, y se agruparon en cuatro grandes

categorías: 1) resultados e impacto del aprendizaje; 2) recursos humanos y financieros invertidos en educación; 3) acceso a la educación, participación y progreso; y 4) ambiente de aprendizaje y organización de las escuelas (OCDE, 2009).

El sistema de indicadores de Canadá se ha desarrollado a partir de la propuesta de la OCDE. Aunque ese país ha participado desde el inicio en *Education at a Glance*, y posee un sistema de indicadores alineado al de la OCDE, su gobierno, preocupado por ofrecer a los políticos y canadienses un sistema que permita monitorear el desempeño del sistema educativo al interior de las provincias, se dio a la tarea en 1996 de desarrollar un sistema de indicadores educativos, que, con el transcurso de los años, fueron modificándose para agruparse en cinco categorías en su edición de 2005 y 2007: 1) características de la población escolar; b) financiamiento del sistema educativo; c) educación elemental secundaria; d) educación postsecundaria; y e) transiciones y resultados. En 2009, las categorías se redujeron a las siguientes: a) los resultados de las instituciones educativas y el impacto del aprendizaje; b) recursos financieros y humanos invertidos en educación; y c) acceso a la educación, participación y progresión. Actualmente, el sistema cuenta con diez indicadores agrupados en las categorías ya mencionadas (Canadian Education Statistics Council, 2005, 2007, 2009).

En España se comenzó a diseñar el sistema estatal de indicadores de la educación a partir de 1993, con la creación del Instituto Nacional de Calidad y Evaluación (INCE), y se publicó su primera versión en 2000. Actualmente, se efectúa una revisión y adecuación anual de 15 indicadores del sistema que han sido denominados prioritarios. El sistema vigente consta de 38 indicadores globales agrupados en cinco categorías: 1) contexto; 2) recursos; 3) escolarización; 4) procesos educativos; y 5) resultados educativos (Ministerio de Educación, Secretaría de Estado de Educación y Formación Profesional, Dirección General de Evaluación y Cooperación Territorial e Instituto de Evaluación, 2009).

En el contexto mexicano, el Instituto Nacional para la Evaluación de la Educación (INEE) ha desarrollado desde 2003 un sistema de indicadores del sistema educativo nacional. En su edición 2008, éste se integró de 37 indicadores, que, a su vez, se subdividen en 297 subindicadores agrupados alrededor de cinco categorías: a) contexto social; b) agentes y recursos en el sistema; c) procesos educativos y gestión; d) acceso y trayectoria; y e) resultados educativos (INEE, 2009).

Otro aspecto con especial relevancia para formar un sistema de indicadores es el establecimiento de los criterios de selección de los indicadores que se incluirán en el sistema. Algunos de los principales criterios que se deben considerar son: la calidad técnica de los datos; su confiabilidad y validez; que los indicadores se deriven de un modelo del sistema educativo y sean políticamente relevantes.

La confiabilidad tiene que ver con la calidad de las medidas e implica el supuesto básico de que los datos serán obtenidos en cada ocasión a lo largo de repetidas observaciones del mismo fenómeno. Por su parte, la validez se refiere a si lo que se está midiendo refleja de manera precisa el concepto que se pretende medir.

Es necesario señalar que la validez y la confiabilidad por sí mismas no cualifican a los indicadores para ser incluidos dentro de un sistema de indicadores. Pueden existir algunos altamente válidos y confiables y, sin embargo, carecer de significado; lo anterior sucede debido a que los indicadores se derivan de un modelo o marco de referencia en particular. Por tanto, si el marco o el modelo carecen de relevancia y significado, los indicadores que se desprendan de él tendrán las mismas carencias, aun cuando sean válidos y confiables.

En este sentido, otro de los criterios que necesariamente determina la selección de indicadores es que éstos se sustenten en un modelo o marco de referencia teórico. Lo anterior implica enfrentar una serie de dificultades, la más importante: poder modelar, de forma adecuada, un fenómeno complejo como es el sistema educativo a partir de un número reducido de elementos.

Uno de los principales criterios para seleccionar indicadores es que éstos sean políticamente relevantes, lo cual permitirá que el sistema de indicadores incluya aspectos cruciales, tanto para las políticas como para los programas. En este mismo sentido, es de suma importancia que los indicadores produzcan información útil en los ámbitos local, estatal y nacional; es decir, deben respetar la autonomía y, al mismo tiempo, ser sensibles para discernir entre definiciones locales y nacionales.

Por último, de acuerdo con lo que plantean Ogawa y Collom (1998), estos criterios no constituyen una receta que deberá seguirse al pie de la letra; son tan sólo una guía que puede orientar el proceso de selección de indicadores y deben, por tanto, utilizarse en concordancia con esta idea.

Aunque el diseño y uso de los indicadores cuenta con una amplia aceptación entre teóricos, investigadores, prácticos y políticos de la educación, este enfoque no está exento de críticas. Para Lashway (2001), por ejemplo, aunque los indicadores encierran la promesa de un mejoramiento del proceso de toma de decisiones, pueden con facilidad provocar confusión. Un peligro latente es la recolección indiscriminada de datos; esto no sólo involucra dinero y esfuerzo, sino que puede atrapar a quienes toman las decisiones en un mar de números, lo cual hace difícil que puedan distinguir lo importante de lo trivial.

Darling-Hammond y Ascher (1991) plantean que las cifras nunca hablan por sí mismas, sino que necesitan una cuidadosa interpretación. Por ejemplo, cuando se observan incrementos en resultados de logro educativo, no se puede de inmediato concluir que se deben a una mejora de la instrucción, ya que existen factores que pudieron influir. Un caso es que los educadores, enfrentados con la necesidad de mejorar las cifras, pudieran tentarse a reemplazar contenidos relevantes con actividades orientadas a la preparación para rendir las pruebas; podrían también excluir a los estudiantes con necesidades educativas especiales en los momentos de las evaluaciones; o incluso podrían cometer fraude.

Darling Hammond y Ascher (1991) hacen notar que los indicadores simplemente proveen información para el sistema, pero que no importa cuán sofisticados sean los datos recolectados, nunca podrán sustituir al juicio humano informado.

COMENTARIOS FINALES

Es necesario señalar que el proceso de selección de indicadores es una tarea influida por múltiples factores, que incluyen, por una parte, las disposiciones legales (leyes, reglamentos, normativas, etcétera) y las metas nacionales e internacionales establecidas para el sistema educativo, y por la otra, los problemas y los retos que la sociedad en su conjunto plantea al sistema educativo, así como la información que proporcionan los resultados de la investigación en el campo educativo.

Aun cuando todos los tipos de indicadores pueden resultar relevantes para la toma de decisiones en materia educativa, son los indicadores de proceso los que pueden revestir una importancia mayor, en primer lugar porque permiten observar lo que sucede en la "caja negra" de la escolarización. Asimismo, tal como señalan Scherens, Glass y Thomas (2005), los indicadores de proceso son interesantes

desde el punto de vista de la política, la gestión y la administración, ya que se refieren a condiciones que son maleables y éstas pueden convertirse en el tema de políticas activas para mejorar la educación.

Los autores opinan que los indicadores de proceso deberían seleccionarse en función de que demuestran asociaciones positivas entre su instrumentación y los resultados educativos. Lo ideal sería que los indicadores del proceso fueran capaces de predecir los resultados, y considerarlos como “funciones de producción de la educación”; es decir, como instrumentos en “proceso” o en condiciones de predecir los incrementos de los resultados de acuerdo con una función exacta.

La necesidad de abundar en las perspectivas de investigación sobre la eficacia escolar y los estudios sobre las buenas prácticas constituye también un argumento significativo para insistir respecto de la identificación y selección de indicadores de proceso.

Asimismo, resulta de suma importancia que los indicadores se encuentren sólidamente sustentados en los resultados más relevantes producidos por la investigación educativa; por ejemplo, los que se presentan en documentos como: “Factores asociados al logro cognitivo de los estudiantes de América Latina y el Caribe” (Treviño *et al.*, 2007), o “Educación de calidad para todos: un asunto de derechos humanos” (OREALC, 2007). Estos documentos contienen información relevante y lineamientos de política educativa claros para la región latinoamericana que se basan en el derecho a la educación establecido por la UNESCO. A partir de la información contenida en ellos es posible determinar nuevos indicadores que permitan el monitoreo y la evaluación de los sistemas educativos latinoamericanos, incluyendo el de México. Cabe señalar que los indicadores utilizados hasta el momento en nuestro país no han sido sensibles a resultados como los mencionados por Hirsch (1999), quien afirma que “... pequeños incrementos en el aprendizaje temprano de la lengua pueden producir enormes consecuencias más adelante” (p. 146). El autor plantea que los niños en edad preescolar que llegan a la escuela con un vocabulario muy estrecho, y por tanto con una base de conocimientos muy limitada y este déficit no se compensa pronto, es casi imposible que logren las habilidades requeridas en grados posteriores, a pesar de que se inviertan esfuerzos importantes en programas de remedio. Las habilidades orales constituyen el fundamento de las habilidades de lectura y escritura. Si el vocabulario oral-sonoro y la comprensión de este vocabulario no se desarrollan de forma correcta y amplia durante etapas tempranas previas a la escolarización, tampoco se desarrollarán las habilidades de lectura y escritura de modo apropiado.

Lo anterior tiene implicaciones para el desarrollo de indicadores y el diseño de políticas educativas que aseguren que los programas preescolares sean académicamente efectivos. Desde nuestra perspectiva, la evaluación de la educación deberá reflejar en sus modelos y sistemas lo más avanzado de las teorías y políticas sociales. Los indicadores son un aspecto que debe definirse de manera consensuada a partir del perfil del ciudadano que se espera lograr como resultado del tránsito de los alumnos a través del sistema educativo nacional. Con base en esto, podrán ponerse en práctica estrategias de monitoreo que permitan valorar el avance hacia el logro del perfil, así como corregir deficiencias y reestablecer el rumbo, e instrumentar los recursos que faciliten este proceso.

En el contexto de esta discusión resulta importante señalar que la definición de indicadores y el establecimiento de estándares y criterios de evaluación son actividades que no sólo deben estar sustentadas sólidamente en las teorías y prácticas

educativas más avanzadas. Deberá, asimismo, reconocerse que no sólo sirven a la función de determinar si se han logrado las metas del sistema, sino que pueden indicar o guiar el rumbo que ha de seguir el propio sistema educativo; en ese sentido, resulta fundamental reconocer el carácter político y ético de los indicadores, estándares y criterios de evaluación.

La selección de un modelo de evaluación refleja no sólo una posición teórica, sino política y ética. Las decisiones respecto de quiénes participan en el diseño de los indicadores, estándares y criterios, la jerarquización de los indicadores, el número de indicadores y el tipo de estándares reflejan la visión de quienes tienen la posibilidad de influir desde la planificación y, en particular, durante el monitoreo del sistema en dirigir y corregir el rumbo de las acciones del sistema educativo, de tal manera que se alcancen los fines que se pretenden en materia educativa.

La solidez de las decisiones dependerá de qué tanto se sustenten éstas en el conocimiento informado respecto de los avances más sobresalientes en materia de teoría y práctica de la evaluación educativa, así como en una discusión profunda sobre las urgentes acciones que se requiere tomar para redirigir las acciones del sistema hacia el logro de metas que revertían los efectos de la exclusión y desigualdad sociales.

Tal como ha señalado la CEPAL/EuropeAid (2007), la construcción de una sociedad democrática en la que se reconozcan y existan los mecanismos para hacer exigibles los derechos básicos, requiere una mayor equidad en las oportunidades educacionales. La aplicación de estrategias adecuadas para elevar los niveles educativos, en especial de los grupos más excluidos (como consecuencia de su negación y devaluación), resultan esenciales para incrementar la movilidad social, elevar la productividad, los retornos entre generaciones y dotarlos de las herramientas necesarias para su integración a la vida moderna, la elevación de su conciencia y participación ciudadana y para la creación de un orden social más justo y meritocrático.

Los sistemas de evaluación pueden contribuir a estos fines apoyando los sistemas educativos en la definición más clara de sus metas y en el establecimiento de sistemas de información que permitan monitorear adecuadamente el sistema y evaluar su impacto en el corto, mediano y largo plazos.

REFERENCIAS BIBLIOGRÁFICAS

- Alkin, M. C. (Ed.) (2004). *Evaluation roots: tracing theorists' views and influences*. Thousand Oaks, CA: Sage publications. Disponible en: http://books.google.com.mx/books?id=swDMkRzVCRIC&dq=Evaluation+roots:+tracing+theorists%2%B4+views+and+influences&printsec=frontcover&source=bn&hl=es&ei=qNKGTOmOG5K4sAO53u2JCG&sa=X&oi=book_result&ct=result&resnum=4&ved=0CCwQ6AEwAw#v=onepage&q&f=false
- Birkmayer, J. y Weiss, C. H. (2000). Theory-Based Evaluation in Practice. What do we learn, *Evaluation Review*, 24 (4), 407–31.
- Blank, R. K. (1993). Developing a system of education indicators: selecting, implementing and reporting indicators. *Educational Evaluation and Policy Analysis*, 11 (2), 65–80. Disponible en: <http://www.eric.ed.gov/PDFS/ED362536.pdf>
- Blank, R. K., y Gruebel, K. (1993). *State indicators of science and mathematics education 1993*. Washington, DC: Council of Chief State School Officers. Disponible en: <http://www.eric.ed.gov/PDFS/ED361225.pdf>
- Canadian Education Statistics Council (2005). *Education indicator in Canada: an international perspective 2005*. Ottawa: Canadian Education Statistics Council. Disponi-

- ble en: <http://www.statcan.gc.ca/pub/81-582-x/2006001/pdf/4225856-eng.pdf>
- _____. (2007). *Report of the Pan-Canadian Education Indicators Program 2007*. Ottawa: Canadian Education Statistics Council. Disponible en: <http://www.statcan.gc.ca/pub/81-582-x/81-582-x2007001-eng.pdf>
- _____. (2009). *Education indicator in Canada: an international perspective 2009*. Ottawa: Canadian Education Statistics Council. Disponible en: <http://www.cmec.ca/Publications/Lists/Publications/Attachments/212/education-indicators-canada-international-perspective-2009.pdf>
- Castells, M. (2002). *La era de la información. Vol. I: La sociedad red*. México, Distrito Federal: Siglo XXI Editores. Disponible en: http://books.google.com.mx/books?id=uADgO-FONJgC&printsec=frontcover&dq=La+Era+de+la+Informaci%C3%B3n+vol.+1&source=gbs_similarbooks_s&cad=1#v=onepage&q&f=false
- CEPAL/EuropeAid (2007). *Un sistema de indicadores para el seguimiento de la cohesión social en América Latina*. Santiago de Chile: Naciones Unidas.
- CERI (1994). *Making education count. Developing and using educational indicators*. París: OCDE.
- Chen, H. T. (2005). *Practical program evaluation. Assessing and improving planning, implementation, and effectiveness*. Newbury Park, CA: Sage. Disponible en: http://books.google.com.mx/books?id=onxLq17MajwC&printsec=frontcover&dq=Practical+program+evaluation&source=bl&ots=KllefmlKgM3&sig=C1C6yAvBCUUAOvGe7z0QEaeqDc&hl=es&ei=kb6GTNzhG12gsQOrmnCDc&sa=X&oi=book_result&ct=result&resnum=1&ved=0CBgQ6AEwAA#v=onepage&q&f=false
- Corvalán, A. M. (2000). *El uso de indicadores: requisito fundamental para alcanzar la educación en el año 2002*. Disponible en: <http://www.unesco.cl/pdf/actyeven/ppe/boletin/artesp/46-1.pdf>
- Dale, R. (1998). *Evaluation Frameworks for Development: Programmes and Projects*. New Delhi: Sage Publications.
- Darling-Hammond, L. y Ascher, C. (1991). *Creating Accountability in Big City School Systems*. Urban Diversity Series No. 102. Nueva York: ERIC Clearinghouse on Urban Education. Disponible en: <http://www.eric.ed.gov/PDFS/ED334339.pdf>
- Dickson, G. S. y Lim, S. (1991). *The development and use of indicators of performance in education leadership*. Ponencia presentada en la Annual Meeting of the International Congress for School Effectiveness and Improvement, Cardiff, Wales, Reino Unido. Disponible en: <http://www.eric.ed.gov/PDFS/ED367044.pdf>
- Donaldson, S. I. y Scriven, M. (2003). *Evaluating social programs and problems: Visions for the new millennium*. Mahwah, NJ: Erlbaum. Disponible en: http://books.google.com.mx/books?id=wd87nB1PZRoC&printsec=frontcover&dq=Evaluating+social+programs+and+problems:+Visions+for+the+new+millennium&hl=es&ei=Ts-GTIHZB4nEsAP49-nDCg&sa=X&oi=book_result&ct=result&resnum=1&ved=0CCcQ6AEwAA#v=onepage&q&f=false
- Escudero, T. (2003). Desde los tests hasta la investigación evaluativa actual. Un siglo, el XX, de intenso desarrollo de la evaluación en educación. *Relieve* 9 (1), 11-43. Disponible en: http://www.uv.es/RELIEVE/v9n1/RELIEVEv9n1_1.htm
- García-Cabrero, B. (1995). La educación preescolar y sus efectos en la prevención de problemas escolares durante la primaria, en G. Acle (ed.). *Educación especial, evaluación, investigación*. México: FES-Zaragoza, UNAM.
- _____. (1986). Evaluación de un programa de educación preescolar con orientación cognoscitiva. Tesis de maestría inédita. Facultad de Psicología, UNAM.

- Hafner, A. y Buchanan, A. (1992). *Education indicators as information Tools: a source-book for school and district staff*. California: The Southwest Regional Educational Laboratory. Disponible en: <http://www.eric.ed.gov/PDFS/ED396359.pdf>
- Hansen, H. F. (2005). Choosing Evaluation Models. A Discussion on Evaluation Design. *Evaluation*, 11 (4), 447-462.
- Hirsch, E. D. (1999). *The schools we need and why we don't have them*. Nueva York: Anchor Books.
- House, E. R. (1993). *Professional evaluation: social impact and political consequences*. Newbury Park: Sage Publications.
- House, E. R. y Howe, K. R. (2001). *Valores en evaluación e investigación social*. Madrid: Morata. Disponible en: http://books.google.com.mx/books?id=rRom4FwTViIC&dq=Valores+en+evaluaci%C3%B3n+e+investigaci%C3%B3n+social&printsec=frontcover&source=bn&hl=es&ei=otaGTP61A4KasAP2p4HeCg&sa=X&oi=book_result&ct=result&resnum=4&ved=0CB8Q6AEwAw#v=onepage&q&f=false
- Instituto Nacional para la Evaluación de la Educación (INEE) (2008). *Panorama educativo de México 2008. Indicadores del sistema educativo nacional*. México: INEE. Disponible en: <http://www.oei.es/pdf2/panorama-educativo-mexico-2008.pdf>
- _____. (2007). *Panorama educativo de México 2007. Indicadores del sistema educativo nacional*. México: INEE. Disponible en: http://www.oei.es/pdfs/panorama_educativo_mexico2007.pdf
- _____. (2006). *Panorama educativo de México 2006. Indicadores del sistema educativo nacional*. México: INEE. Disponible en: <http://www.oei.es/quipu/mexico/01-panoramaweb.pdf>
- _____. (2005). *Panorama educativo de México 2005. Indicadores del sistema educativo nacional*. México: INEE. Disponible en: <http://www.inee.edu.mx/index.php/publicaciones/informes-institucionales/panorama-educativo/3414>
- _____. (2004). *Panorama educativo de México 2004. Indicadores del sistema educativo nacional*. México: INEE. Disponible en: <http://www.inee.edu.mx/index.php/publicaciones/informes-institucionales/panorama-educativo/3415>
- _____. (2003). *Panorama educativo de México 2003. Indicadores del sistema educativo nacional*. México: INEE. Disponible en: <http://www.inee.edu.mx/index.php/publicaciones/informes-institucionales/panorama-educativo/3416>
- Joint Committee on Standards for Educational Evaluation (1994). *The Program Evaluation Standards*, 2a edición. Newbury Park, CA: Sage Publications. Disponible en: <http://www.jcsee.org/program-evaluation-standards>
- Jones, R. M. y Nielsen, J. I. (1994). Saskatchewan's education indicators program. *Education Canada*, 4-8.
- Joo, K. P. (2009). Monitoring as a Program Evaluation Model: A Case Study of Developing Evaluation Criteria for Monitoring in a Korean University Lifelong Education Center. Ponencia presentada en The Midwest Research-to-Practice Conference in Adult, Continuing, Community and Extension Education, Northeastern Illinois University, Chicago, IL, 21 al 23 de octubre. Disponible en: <http://www.neiu.edu/~hrd/mwr2p09/Papers/Joo.pdf>
- Madaus, G. F. y Kellaghan, T. (2000). Models, metaphores and definitions in evaluation, en D. Stufflebeam, G. F. Madaus y T. Kellaghan. *Evaluation models: viewpoints on educational and human services* (19-32). Boston: Kluwer Academic Publishers. Disponible en: <http://books.google.com.mx/books?id=5679vdyleZMC&printsec=frontcover&dq=Evaluation+models:+viewpoints+on+educational+and+human+>

- ervices&hl=es&ei=f82GTObaOI7EsAPWnuidCg&sa=X&oi=book_result&ct=result&resnum=1&ved=0CCcQ6AEwAA#v=onepage&q&f=false
- Mateo, J. (2006). *La evaluación educativa, su práctica y otras metáforas*. México: Alfaomega Grupo Editor.
- Ministerio de Educación/Secretaría de Estado de Educación y Formación Profesional/ Dirección General de Evaluación y Cooperación Territorial/ Instituto de Evaluación (2009). *Sistema Estatal de Indicadores de la Educación*. Madrid: Ediciones Gráficas Arial S.L. Disponible en: <http://www.ince.mec.es/pub/index.htm>
- Morduchowicz, A. (2006). *Los indicadores educativos y las dimensiones que los integran*. Buenos Aires: LIPE-UNESCO. Disponible en: http://148.202.105.241/biblioteca/bitstream/123456789/2134/1/Sistemas+de+indicadores_educativos+Argentina.pdf
- Muñoz-Izquierdo, C. y Márquez, A. (2000). Indicadores del desarrollo educativo en América Latina y de su impacto en los niveles de vida de la población. *Revista Electrónica de Investigación Educativa*, 2 (2). Disponible en: http://redie.uabc.mx/vol2no2/contents-munoz.html/modules/Forums/admin/admin_users.php
- OECD (1998). *Education at glance*. OECD Publishing. Disponible en: http://www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/15/b6/82.pdf
- ____ (2009). *Education at glance*. OECD Publishing. Disponible en: <http://www.oecd.org/dataoecd/41/25/43636332.pdf>
- Ogawa, R. y Collom, E. (1998). *Educational indicators: What are they? How Can Schools and Schools Districts Use Them?*. California: California Educational Research Cooperative. Disponible en: <http://www.eric.ed.gov/PDFS/ED432811.pdf>
- Pawson, R. (2002). Evidence-Based Policy: The Promise of Realist Synthesis. *Evaluation*, 8 (3), 340–58.
- Pawson, R. y Tilley, N. (1997). *Realistic Evaluation*. Londres: SAGE. Disponible en: http://books.google.com.mx/books?id=Oudpyutx8_QC&printsec=frontcover&dq=Realistic+Evaluation&source=bl&ots=sK8NYp9QUP&sig=8pSqVfxtkG_7NXMJ33VadLKN_FA&hl=es&ei=HRGHITZLJO6sQPpy3vS6Cg&sa=X&oi=book_result&ct=result&resnum=1&ved=0CBkQ6AEwAA#v=onepage&q&f=false
- Payne, D. A. (1994). *Designing educational project and program evaluations: A practical overview based on research and experience*. Boston: Kluwer Academic Publishers. Disponible en: http://books.google.com.mx/books?id=7a8_rLxHbcC&printsec=frontcover&dq=Designing+educational+project+and+program+evaluations&hl=es&ei=ZRGHTIL9OJS6sQPxq8DMCg&sa=X&oi=book_result&ct=result&resnum=1&ved=0CC4Q6AEwAA#v=onepage&q&f=false
- Programa Sectorial de Educación 2007-2012 (2007). México: SEP. Disponible en: http://upepe.sep.gob.mx/prog_sec.pdf
- Rossi, Peter H., Lipsey, M. W. y Freeman, E. H. (2004). *Evaluation: A Systematic Approach*. Thousand Oaks, CA: SAGE. Disponible en: http://books.google.com.mx/books?id=8K3Ax5s7x5oC&printsec=frontcover&dq=Evaluation:+A+Systematic+Approach&hl=es&ei=usKGTNDmDIm4sQOxtPS2Cg&sa=X&oi=book_result&ct=result&resnum=1&ved=0CDEQ6AEwAA#v=onepage&q&f=false
- Scheerens, J. y Creemers, B. P. (1989). Towards a more comprehensive conceptualization of school effectiveness, en B. P. Creemers, T. Peters y D. Reynolds (eds.). *School effectiveness and school improvement* (265-278). Lisse: Swets & Zeitlinger.
- Scheerens, J., Glass, C. y Thomas, S. (2005). *Educational Evaluation, Assessment, and*

- monitoring. A Systemic Approach*. Taylor & Francis e-Library.
- Schwandt, T. A. (2005). The Centrality of Practice to Evaluation. *American Journal of Evaluation*, 26 (1), 95-105.
- Scott, W. R. (2003) *Organizations: Rational, Natural and Open Systems*. Englewood Cliffs, J: Prentice-Hall.
- Shavelson, R. J., McDonnell, L. M. y Oakes, J. (eds.) (1989). *Indicators for Monitoring Mathematics and Science Education: A sourcebook*. Santa Monica: RAND Corporation.
- Shepard, L., Hannaway, J. y Baker, E. (2009). *Standards, Assessment and Accountability*. Education Policy White Paper. Disponible en: <http://www.naeducation.org>
- Stainmetz, (1983). The discrepancy evaluation model, en G. F. Madaus, M. Scriven y D. L. Stufflebeam (eds.). *Evaluation Models: Viewpoints on Educational and Human Services Evaluation* (79-100). Boston: Kluwer-Nijhoff. Disponible en: http://books.google.com.mx/books?id=5679vdyleZMC&printsec=frontcover&dq=Evaluation+Models:+Viewpoints+on+Educational+and+Human+Services+Evaluation&hl=es&ei=jRSHTJP1HoGglAfH7aDeCA&sa=X&oi=book_result&ct=result&resnum=1&ved=0CCoQ6AEwAA#v=onepage&q&f=false
- Stake, R. (1967.). *The countenance of educational evaluation*. *Teachers College Record*, 68, 523-540.
- ____ (2006). *Evaluación comprensiva y evaluación basada en estándares*. Barcelona: Editorial Graó.
- Strater, D. (2006). *Standard-based assessment*. Disponible en: <http://filebox.vt.edu/users/dstrater/StandardsBasedAssessment.pdf>
- Stufflebeam, D. L. (2000). The CIPP model for evaluation, en D. L. Stufflebeam, G. F. Madaus y T. Kellaghan (eds.). *Evaluation models. Viewpoints on educational human services* (279-317). Boston: Kluwer Academic Publishers. Disponible en: http://books.google.com.mx/books?id=5679vdyleZMC&printsec=frontcover&dq=Evaluation+models.+Viewpoints+on+educational+human+services&hl=es&ei=QMSGTIa0D4T0tgPA7ODTCg&sa=X&oi=book_result&ct=result&resnum=1&ved=0CCoQ6AEwAA#v=onepage&q&f=false
- ____ (2004). The 21st-Century CIPP Model: Origins, Development, and Use, en M. C. Alkin (ed.). *Evaluation roots* (pp. 245-266). Thousand Oaks, CA: Sage publications. Disponible en: http://books.google.com.mx/books?id=X2vsczE0CXwC&printsec=frontcover&dq=Evaluation+roots&hl=es&ei=fxWHTKHqAYScIgf26JX2Dw&sa=X&oi=book_result&ct=result&resnum=1&ved=0CCoQ6AEwAA#v=onepage&q&f=false
- ____ (1968). *Evaluation as enlightenment for decision-making*. Columbus, Ohio: Evaluation Center, Ohio State University.
- Stufflebeam, D. y Shinkfield, A. J. (1987). *Evaluación sistemática: guía teórica y práctica*. Madrid: Paidós-MEC. Barcelona: Paidós.
- Tiana, A. (2008). Evaluación y cambio de los sistemas educativos: la interacción que hace falta. *Revista Iberoamericana de Educación* 16 (59), abril-junio, Rio de Janeiro. Disponible en: <http://www.oei.es/noticias/spip.php?article3840>
- Tognolini, J. y Stanley, G. (2007). Standards-based assessment: a tool and means to the development of human capital and capacity building in education. *Australian Journal of Education*, 51(12), 129-145.
- Vedung, E. (1997). *Public Policy and Program Evaluation*. New Brunswick, NJ: Transaction Publishers. Disponible en: http://books.google.com.mx/books?id=VbUN9cOHGIC&printsec=frontcover&dq=Public+Policy+and+Program+Evaluation&hl=es&ei=wxWHTOj-BsaAlAFqVgFdw&sa=X&oi=book_result&ct=result&resnum

=1&ved=0CC4Q6AEwAA#v=onepage&q&f=false

Wholey, J. S. (2004). Using Evaluation to Improve Performance and Support Policy Decisionmaking, en M. C. Alkin. *Evaluation roots* (267-275). Thousand Oaks, CA: Sage. Disponible en: http://books.google.com.mx/books?id=X2vsczE0CXwC&printsec=frontcover&dq=Evaluation+roots&hl=es&ei=fxWHTKHqAYScIgf26JX2Dw&sa=X&oi=book_result&ct=result&resnum=1&ved=0CCoQ6AEwAA#v=onepage&q&f=false