

MODELOS TEÓRICOS E INDICADORES DE EVALUACIÓN EDUCATIVA

BENILDE GARCÍA CABRERO

Av. Universidad 3004, Col. Copilco-Universidad, Coyoacán, C.P. 04510
México, D.F.

Tel. 56 22 22 22 Ext. 41231

Correo electrónico: benilde@servidor.unam.mx

Profesora Titular del Departamento de Psicología Educativa y Desarrollo de la División de Estudios de Posgrado de la Facultad de Psicología de la UNAM en la que imparte cursos en los niveles de licenciatura, maestría y doctorado. Miembro del Sistema Nacional de Investigadores de Conacyt y Consejera Técnica del Instituto Nacional para la Evaluación de la Educación (INEE), donde colabora con la Dirección de Indicadores Educativos. Profesora de diversas universidades nacionales y evaluadora de programas en instituciones nacionales e internacionales, entre las que se encuentran: UNICEF, UNESCO, OCDE, CONAFE, IFE y SEP, entre otras.

RESUMEN

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

PALABRAS CLAVE:

Monitoreo de la educación, modelos de evaluación educativa, indicadores, estándares.

INTRODUCCIÓN

El conocimiento y la información constituyen de acuerdo con diversos expertos, los elementos centrales de las diferencias entre las personas, las instituciones y los países en la época actual (Castells, 2002). De acuerdo con lo planteado por Mateo (2006), el objetivo fundamental de los esfuerzos de desarrollo de las sociedades industriales se orientó hacia el dominio técnico de la naturaleza, en tanto que las sociedades postindustriales que existen en la época actual se orientan fundamentalmente hacia la generación y la gestión democrática del conocimiento. En este contexto, para que el desarrollo económico se sustente, se requiere invertir en capital humano, en elevar los conocimientos de los ciudadanos y las habilidades de los mismos para sacar el mayor provecho de la información disponible. 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 elites. El dominio de la tecnología en 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 y /o comercializan, produciendo con ello situaciones de injusticia y de marginación que atentan contra el bienestar y la cohesión social.

Ante este panorama, Mateo (2006) señala que la articulación de mecanismos socioeducativos que permitan que cualquier persona en cualquier etapa de su vida pueda tener acceso a desarrollar instrumentos y competencias cognitivas que le permitan 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"(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, sus procesos, sus agentes y sus resultados englobando en la medida de lo posible al conjunto del sistema educativo.

Para responder a las necesidades y demandas actuales en materia educativa, así como a los requerimientos de la globalización y de la sociedad del conocimiento, el Programa Sectorial de Educación 2007-2012, propuso impulsar una profunda reforma educativa que se encuentra actualmente en marcha. Como parte de esta reforma se ha planteado la necesidad de

promover la evaluación de todos los actores y procesos involucrados en el sistema educativo. Para ello es necesario monitorear 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, se requiere 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 del sistema 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 análisis de información –de forma rutinaria 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, preferiblemente para ajustar la etapa de planificación del ciclo y proponer acciones correctivas.

Por su parte, la evaluación se define como un examen más profundo, que se realiza en determinados momentos de la operación de los programas o proyectos o de partes de ellos, por lo general con énfasis en el impacto, la eficiencia, eficacia, pertinencia, replicabilidad y sostenibilidad (Dale, 1998).

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 fundamentalmente en datos administrativos; en la realización de esta actividad evaluadora 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 monitoreo del proceso, como al monitoreo de los resultados (Chen, 2005; Fitzpatrick, Rossi, Lipsey y Freeman, 2004; Sanders y Whorthen, 2004). Joo (2009) señala que el monitoreo del proceso es de gran ayuda para verificar que la implementación del programa se esté realizando conforme a lo planeado.

El monitoreo de los resultados se puede realizar a través de diferentes medios, particularmente 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. Una de las estrategias que ha sido adoptada por numerosos países, particularmente los más avanzados, es la 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 función de resultados predeterminados para las diferentes asignaturas del curriculum.

Un componente central para el monitoreo de los sistemas educativos son los estándares, que pueden ser definidos como criterios claros y públicos que establecen las referencias de lo que los alumnos pueden y deben saber y

saber hacer, en cada una de las asignaturas de los planes de estudio correspondiente 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 lograrla. 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 en general puedan solicitar a las escuelas y al sistema educativo en general, 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. Para determinar si los estudiantes logran los estándares, es necesario diseñar tareas de evaluación o exámenes acordes con dichos estándares. Los estándares no están referidos a una norma y los reportes acerca del avance de los estudiantes no se refieren 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.

El sistema de indicadores educativos que publica la OCDE desde 1998 bajo el título de *Education at a Glance*, constituye un ejemplo de un sistema de monitoreo que pretende establecer comparaciones entre diferentes países en relación con la situación del sistema educativo.

Education at a Glance incluye las visiones más actuales sobre la construcción y 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 evolución de los sistemas de evaluación y monitoreo en diversos países a nivel mundial.

En el contexto nacional, la publicación anual del Instituto Nacional para la Evaluación de la Educación (INEE) denominada Panorama Educativo desde 2003 que se ha orientado básicamente a monitorear/dar seguimiento a la calidad de la educación que se ofrece en nuestro país, constituye un ejemplo del interés creciente por la evaluación. El INEE concibe la calidad como la coherencia de las relaciones entre componentes orientada al logro sostenido de los objetivos educativos para contribuir a la superación de las necesidades sociales. Para dar cuenta de la calidad, la Dirección de Indicadores del INEE ha diseñado y aplicado un sistema de monitoreo de indicadores de contexto, insumo, proceso y producto.

En el contexto de esta discusión, cabe señalar que el monitoreo de las características del sistema educativo puede realizarse considerando diferentes niveles de agregación: el sistema educativo, un programa específico, una escuela, un salón de clases o una cohorte determinada de alumnos.

Los términos “criterio de evaluación” y “estándar de evaluación” a menudo son confundidos. 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 a dos cosas: por una parte, el criterio (en el sentido que acabamos de definir en líneas anteriores), y una norma sobre la base de la cual se puede decidir 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, mismos que 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 5^o. grado se propone que los niños aprendan a redactar informes. Sin embargo, no basta con señalar que se espera que los niños de educación básica sean capaces de redactar informes, sino que se requiere señalar por ejemplo, cuántos informes deberán realizar, o qué características mínimas deben contener éstos. Considerando que los programas de las asignaturas están formulados por bloques, uno por cada uno de los cinco bimestres, el estándar deberá señalar si se espera que se desarrolle un informe por cada bloque, además de especificar los criterios de calidad con que deberán redactarse. Sin este referente, resulta complejo establecer mecanismos de monitoreo preciso de los procesos y resultados del sistema.

La literatura actual sobre el tema de evaluación señala que las diversas aproximaciones que se han desarrollado hasta el presente, coinciden en que su propósito es reunir información sistemática y hacer algún tipo de juicio sobre un determinado objeto de evaluación. En el ámbito educativo, estos objetos de evaluación pueden ser: los alumnos, las escuelas, el tipo de

servicio educativo, el sistema educativo, etc. (Hansen, 2005; Stufflebeam, 2000, Suchman, 1979).

Scheerens, Glass y Thomas (2005) señalan 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, datos que se basan en medidas del rendimiento educativo, los que están disponibles en registros administrativos (incluyendo estadísticas educativas) y aquellos que provienen de la revisión de expertos y la investigación educativa.

De acuerdo con lo señalado por Scheerens, Glass y Thomas (2005), el propósito fundamental de la evaluación del sistema educativo, es la determinación empírica de la calidad de dicho sistema. 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 determinar si las características del objeto evaluado se conforman de manera formal a las normas y estándares establecidos. Por su parte, la rendición de cuentas permite que la calidad del objeto pueda ser inspeccionada por otras instancias de la sociedad en general. 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 vs. interna de los procedimientos de evaluación y en la orientación sumativa vs. formativa que emplean.

En el caso de la certificación y acreditación, la evaluación tiene un alto grado de formalidad, es de tipo sumativa y para desarrollarla se requieren estándares claramente especificados para certificar estudiantes o profesionales y para la acreditación de programas. La rendición de cuentas, por su parte, requiere de una evaluación formal de tipo externa con propósitos de control en la que se combinan los enfoques formativo y sumativo, 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, son algunas de las actividades que han implementado países como el Reino Unido, Holanda, Francia, Argentina, Chile, República Dominicana y México para la mejora de sus respectivos sistemas educativos. Así mismo, diversos organismos internacionales (OCDE, UNESCO, Unión Europea, OEI, etc.) 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 ha implicado transformaciones en la concepción y práctica de la evaluación, particularmente cambios conceptuales en las nociones monolíticas de evaluación, sustituyéndolas 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, desarrollándose una creciente tendencia a la integración de métodos cuantitativos y cualitativos. Por su parte, los usos de la evaluación se han modificado enfatizándose su carácter político y su capacidad como herramienta para seleccionar lo que se pretende evaluar y por tanto en la posibilidad de influir e la orientación de los sistemas educativos.

Se han introducido también cambios estructurales en la evaluación, que se caracterizan por la creciente inclusión de la evaluación en los mecanismos de gestión de los sistemas educativos, así como por una ampliación de sus ámbitos de cobertura y una mayor interdisciplinariedad de la misma (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 producido como efecto, una rápida evolución de la evaluación, concebida como disciplina científica y como práctica profesional (Tiana, 2008).

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 que se utilizan para llevar a cabo la evaluación (Donaldson y Scriven, 2003).

Este supuesto consenso respecto del propósito de la evaluación ha sido discutido por algunos evaluadores, señalando que su responsabilidad fundamental es la descripción y la explicación científica de las relaciones que existen entre indicadores, y que el juicio de valor no es responsabilidad de los evaluadores, sino de los interesados. Esta perspectiva, que ha sido 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 determinarse qué valor tienen. Por tanto, determinar la utilidad (efectividad-resultados) de un programa, en función del logro de los objetivos que se propone, es realmente 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, etc.) 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 hayan tomado en cuenta todos los aspectos que determinan el mérito, valía o significatividad del objeto evaluado, mediante una detección 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 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 respecto de un determinado objeto de evaluación revela, en gran medida la postura adoptada en relación con 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, 1986). En la literatura sobre evaluación educativa existen diversos modelos formales de evaluación cuyas representaciones ayudan a examinar las relaciones entre diferentes componentes y procesos que tienen lugar dentro de dichos sistemas (e. g. Provus, 1972; Stake, 1967, Stufflebeam, 1968).

Madaus y Kellaghan (2000) señalan que algunos autores han criticado la utilización del término modelo para referirse a las diferentes 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 a esta crítica y como defensa de la utilización del término modelo,

los autores señalan que las definiciones comunes de los diccionarios plantean que un modelo es una síntesis o abstracción de un fenómeno o proceso. Si se considera a la evaluación como un proceso, entonces los modelos de evaluación constituyen la forma como un determinado autor resume o conceptualiza la manera como debe realizarse 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 del 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 presupuestal, 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, junto con 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 que existen actualmente 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 señalar 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 se requiere 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 constituyen el sistema (Ogawa y Collom, 1998).

Hansen (2005) realiza 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 (Stufflebeam, 2000). El segundo enfoque consiste en una valoración retrospectiva y cuidadosa del mérito y valor de la administración, productos y resultados de intervenciones gubernamentales dirigidas a jugar un papel en relación con 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) señalan que los modelos básicos de evaluación que se utilizan 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, de acuerdo con Hansen (2005), 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.

El modelo de resultados también denominado modelo basado en las metas, se centra en los productos alcanzados por un determinado programa u organización. Existen dos subcategorías del modelo de resultados: 1) el modelo de logro de metas, y 2) el modelo de efectos. El modelo de logro de metas es un clásico en la literatura sobre evaluación de programas y evaluación organizacional; en el que los resultados se evalúan con base en las metas u objetivos previamente establecidos (Phopham, 1969, Provus, 1972; Scriven, 1973; Stainmetz, 1983; Tyler, 1968). 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), ha sido criticado debido a que 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 que se están llevando a cabo. 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 (Hernichsfeger y Wiley, 1989; Shulman, 1986, Walberg 1976).

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 establecidos 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 basa 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.

Finalmente, 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 este modelo, la evaluación se basa 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 Sistema de Indicadores del Instituto Nacional para la Evaluación de la Educación (INEE, 2003, 2004, 2005, 2006, 2007, 2008)

En las últimas cuatro décadas, particularmente en los 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 estos modelos 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 lo señalado por Hansen (2005), a la decisión de elegir una perspectiva determinada 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 términos del logro académico e impacto social, mismos que a su vez influyen nuevamente en el contexto.

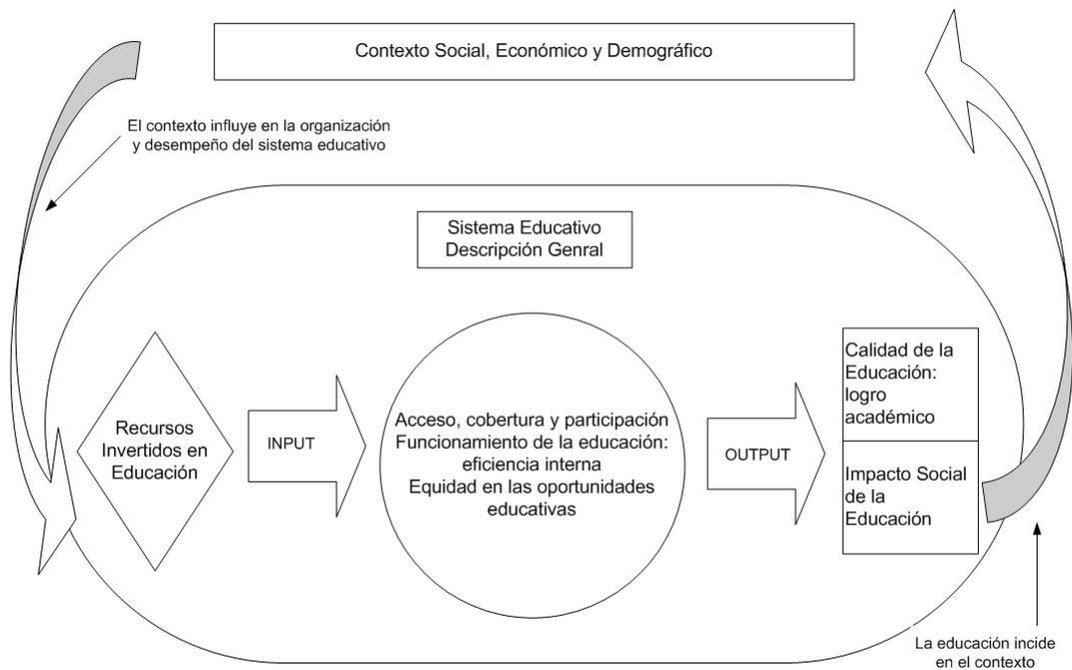


Figura 1. Modelo interpretativo de relaciones. Tomado de Corvalán (2000).

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, 5) Impacto Social de la Educación.

Contexto Demográfico, Social y Económico y Descripción General del Sistema de Educación.

Estos indicadores, de acuerdo con Scheerens, Glass, y Thomas (2005) son generalmente 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 de estos indicadores 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

Son aquellos indicadores que 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, porcentaje de la fuerza laboral del país que se dedica al campo educativo.

Desempeño de los Sistemas Educativos

Los indicadores que se incluyen en esta categoría, dan cuenta de las características del entorno de aprendizaje, así como de los aspectos organizacionales de la escuela. Algunos ejemplos de este tipo de indicadores 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 arreglos estructurales para el monitoreo y 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, pero sobre todo, con los datos relacionados con el logro educativo. Algunos ejemplos de

este tipo de indicadores 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, etc.).

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 determinado nivel de certificación en el mercado laboral, ingresos relacionados con la educación y nivel de entrenamiento de los individuos, entre otros. Un ejemplo del uso de este tipo de indicadores se presenta en el análisis realizado por 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 del número que deberá incluirse en el sistema. 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; Porter, 1991; Shavelson, McDonnell y Okes, 1989). Por consiguiente, es deseable que los sistemas de indicadores contemplen sólo un número esencial de indicadores. No obstante, aun existe una falta de consenso en relación con la cantidad de indicadores que puede considerarse aceptable y funcional dentro un sistema de indicadores en particular. Algunos autores, como Dickson y Lim (1991)

plantean que se deben incluir entre cinco y ocho indicadores; otros autores (e. g. Blank, 1993) señalan que basta con incluir doce indicadores, mientras que algunos como Hafner y Buchanan (1992) recomiendan un máximo de veinte.

Resulta evidente que el número de indicadores a incluir en un sistema, 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 en el sistema, se debe también considerar el contexto particular en el que se inscribe el sistema de indicadores, y en función del mismo, determinar el número de indicadores que conformarán el sistema.

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 (15 a 29 años), así como seguimiento de la eficiencia termina 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). Para el año 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 Canadá ha participado desde el inicio en

Education at a Glance, y posee un sistema de indicadores alineado al de la OCDE, el gobierno de Canadá, preocupado por ofrecer a los políticos y canadienses un sistema que permita monitorear el desempeño del sistema educativo al interior de las provincias del país, se dio a la tarea en el año 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 del año 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 el año 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 10 indicadores agrupados en las categorías mencionadas arriba (*Canadian Education Statistics Council*, 2005, 2007, 2009).

En España se comenzó a diseñar el sistema estatal de indicadores de la educación a partir del año 1993, con la creación del Instituto Nacional de Calidad y Evaluación (INCE), publicándose su primera versión en el año 2000. Actualmente se realiza una revisión y adecuación anual de 15 indicadores del sistema que han sido denominados prioritarios. El sistema de indicadores vigente consta de 38 indicadores globales agrupados en cinco categorías: 1) Contexto, 2) Recursos, 3) Escolarización, 4) Proceso 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 el año 2003, un sistema de indicadores del sistema educativo nacional. En su edición 2008, el sistema de indicadores se conformó por 37 indicadores, que a su vez se subdividen en 297 sub-indicadores que se agrupan 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 que resulta de especial relevancia para conformar 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, la confiabilidad y validez de los mismos, que los indicadores se deriven de un modelo del sistema educativo y que sean políticamente relevantes.

La confiabilidad tiene que ver con la calidad de las medidas e implica el supuesto básico de que los mismos 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 indicadores altamente válidos y confiables y sin embargo carecer de significado alguno, lo anterior sucede debido principalmente 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 del mismo tendrán las mismas carencias, aún 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. Esto permitirá que el sistema de indicadores incluya aspectos cruciales, tanto para las políticas como para los programas. En este mismo sentido, cabe destacar que es de suma importancia que los indicadores produzcan información útil en los niveles local, estatal y nacional; es decir, deben respetar la autonomía y al mismo tiempo ser sensibles para discernir entre definiciones locales y nacionales.

Finalmente, de acuerdo con lo que plantean Ogawa y Collom (1998), es necesario señalar, que 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. Lashway (2001) por ejemplo, señala que aunque los indicadores conllevan la promesa de un mejoramiento del proceso de toma de decisiones, pueden fácilmente 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 de 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 puede haber una serie de factores que pudieron influir en los mismos. Por ejemplo, enfrentados con la necesidad de mejorar las cifras, los educadores 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 que se ve influida por múltiples factores. Estos factores incluyen, por una parte, las disposiciones legales (leyes, reglamentos, normatividades, etc.), y las metas nacionales e internacionales establecidas para el sistema educativo, y por la otra, las problemáticas y 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.

Aún 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 señalan que los indicadores de proceso deberían seleccionarse en función de que demuestran asociaciones positivas entre la su instrumentación y los resultados educativos. Lo ideal sería que los indicadores del proceso fueran capaces de predecir los resultados (como si fueran "funciones de producción de la educación": 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 importante 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 y que se reflejan 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), 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 estos documentos, es posible establecer nuevos indicadores que permitan el monitoreo y la evaluación de los sistemas educativos latinoamericanos, incluyendo el de México.

También cabe señalar que los indicadores que se han utilizado hasta el momento en nuestro país, no han sido sensibles a resultados como los señalados 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 que se requieren 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 apropiadamente.

Lo anterior tiene implicaciones para el desarrollo de indicadores y para 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 concensuada 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 establecerse estrategias de monitoreo que permitan valorar el avance hacia el logro del perfil, así como corregir deficiencias y re-establecer el rumbo, así como instrumentar los recursos que faciliten este proceso.

REFERENCIAS

- Alkin, M. C. (Ed.) (2004). *Evaluation roots: tracing theorists' views and influences*. Thousand Oaks, CA: Sage publications
- Birckmayer, J. y Weiss, C. H. (2000). Theory-Based Evaluation in Practice, *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.
- Blank, R. K., y Gruebel, K. (1993). *State indicators of science and mathematics education 1993*. Washington, D.C.: Council of Chief State School Officers
- Canadian Education Statistics Council (2005). *Education indicator in Canada: an international perspective 2005*. Ottawa: Canadian Education Statistics Council. Disponible en: <http://www.statcan.gc.ca/pub/81-582-x/2006001/pdf/4225856-eng.pdf>
- Canadian Education Statistics Council (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>
- Canadian Education Statistics Council (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.

CERI (1994). *Making education count. Developing and using educational indicators*. París: OCDE

Chen, H. T. (2005). *Practical program evaluation*. Newbury Park, CA: Sage.

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.

Darling-Hammond, L. y Ascher, C. (1991). *Creating Accountability in Big City School Systems*. Urban Diversity Series # 102. New York: ERIC Clearinghouse on Urban Education. ED 334-339.

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.

Donaldson, S.I. y Scriven, M.(2003). *Evaluating social programs and problems: Visions for the new millennium*. Mahwah, NJ: Erlbaum.

- Fitzpatrick, J. L., Sanders, J. R., y Worthen, B. R. (2004). *Program evaluation: Alternatives approaches and practical guides* (3rd ed.). Nueva York: Longman.
- García-Cabrero, B. (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 sourcebook for school and district staff*. California: The Southwest Regional Educational Laboratory.
- Hansen, H.F. (2005). Choosing Evaluation Models. A Discussion on Evaluation Design. *Evaluation*, 11(4), 447-462
- Hernichsfeger, A. y Wiley, D. C. (1978). Conceptual issues in models of school learning. *Curriculum Studies*, 10 (3), 215-231.
- Hirsch, E.D. (1999). *The schools we need and why we don't have them*. New York: anchor Books.
- House, E. R. (1993). *Professional evaluation: social impact and political consequences*. Newbury Park: Sage
- House, E. R. y Howe, K. R. (2001). *Valores en evaluación e investigación social*. Madrid: Morata.
- Instituto Nacional para la Evaluación de la Educación (INEE) (2008). *Panorama Educativo de México 2008. Indicadores del Sistema Educativo Nacional*. México: INEE.

Instituto Nacional para la Evaluación de la Educación (INEE) (2007). *Panorama Educativo de México 2007. Indicadores del Sistema Educativo Nacional*. México: INEE.

Instituto Nacional para la Evaluación de la Educación (INEE) (2006). *Panorama Educativo de México 2006. Indicadores del Sistema Educativo Nacional*. México: INEE.

Instituto Nacional para la Evaluación de la Educación (INEE) (2005). *Panorama Educativo de México 2005. Indicadores del Sistema Educativo Nacional*. México: INEE.

Instituto Nacional para la Evaluación de la Educación (INEE) (2004). *Panorama Educativo de México 2004. Indicadores del Sistema Educativo Nacional*. México: INEE.

Instituto Nacional para la Evaluación de la Educación (INEE) (2003). *Panorama Educativo de México 2003. Indicadores del Sistema Educativo Nacional*. México: INEE.

Joint Committee on Standards for Educational Evaluation. (1994). *The Program Evaluation Standards*, 2a Edición. Newbury Park, CA: Sage Publications.

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.

Lasway, L. (2001). *Educational Indicators*. ERIC Digest 150. Disponible en: <http://eric.uoregon.edu/publications/digests/digest150.html>

Madaus, G. F. y Kellaghan, T. (2000). Models, metaphores and definitions in evaluation. En: G. F. Madaus y T. Kellaghan, *Evaluation models: viewpoints on educational and human services* (19-32). Boston: Kluwer Academic Publishers.

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 e Instituto de Evaluación (2009). *Sistema Estatal de Indicadores de la Educación*. Madrid: Ediciones Gráficas Ariel S.L.

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*. Vol. 2, No.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_01/000019b/80/15/b6/82.pdf
- OECD (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.
- 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.
- Payne, D.A. (1994). *Designing educational project and program evaluations: A practical overview based on research and experience*. Boston: Kluwer Academic Publishers.
- Popham, J.W. (1969). Objectives and instruction. En W.J. Popham, E.W. Eisner, H.J. Sullivan y L.L.
- Porter, A. (1991). Creating a system of school process indicators. *Educational Evaluation and Policy Analysis*, 11(2), 13-29.
- Programa Sectorial de Educación 2007-2012 (2007). México: SEP. Disponible en: http://upepe.sep.gob.mx/prog_sec.pdf

- Provus, M. (1972). The discrepancy evaluation model. En: P. A. Taylor y D. M. Cowley (Eds.), *Readings in curriculum evaluation*. Dubuque, Iowa: Wm. C. Brown Co. Publishers.
- Rossi, Peter H., Lipsey, M. W., Freeman, E. H. (2004). *Evaluation: A Systematic Approach*. Thousand Oaks, CA: SAGE.
- 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.
- Scriven, M. (1973). Goal-Free Evaluation. En Ernest R. House (Ed.) *School Evaluation: The Politics and Process*, pp. 319–28. Berkeley, CA: McCutchan.
- Shavelson, R.J., L.M. McDonnell, J. Oakes (Eds) (1989). *Indicators for Monitoring Mathematics and Science Education: A sourcebook*. Santa Monica: RAND Corporation.
- Sheerens, J. y Creemers, B. P. (1989). Towards a more comprehensive conceptualization of school effectiveness. En: BP. Creemers, T. Peters y D. Reynolds (Eds.), *School effectiveness and school improvement* (265-278). Lisse: Swets & Zeitlinger.

- Shulman, L. S. (1986). Paradigms and research programs in the study of teaching: A contemporary perspective. En: M.C. Wittrock (Ed.), *Handbook of Research on Teaching* (3-36). New York: Macmillan.
- 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.
- Stake, R. (1967.). *The countenance of educational evaluation. Teachers College Record*, 68, 523-540.
- Stake, R. (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. (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.
- 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.

- Stufflebeam, D. L. (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
- Suchman, E. (1979). *Evaluation Research*. Rusell: Sage Foundation
- Tiana, A. (2008). Evaluación y cambio de los sistemas educativos: la interacción que hace falta. *Revista Iberoamericana de Educación*. vol.16 no.59 Rio de Janeiro Apr./June 2008. 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 humen capital and capacity building in education. *Australian Journal of Education*, 51(12), 129-145.
- Tyler, R. W. (1968). *The Challenge Of National Assessment*. Columbus, Oh: Merrill.
- Vedung, E. (1997). *Public Policy and Program Evaluation*. New Brunswick, NJ: Transaction Publishers.
- Walberg, H. J. (1976). Psychology of learning enviroments: Behavioral, structural or perceptual?. *Review of Research on Education*, (4), 142-178.
- 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.