

*Otro concepto central en la discusión es el de **sistema de indicadores**. Sea que se trate de estadísticas simples o compuestas, un indicador singular difícilmente podrá proporcionar información útil sobre fenómenos tan complejos como los educativos. Los sistemas de indicadores se diseñan por lo general para generar información más amplia y precisa. Pero debe precisarse que un sistema de indicadores es más que una simple colección de estadísticas. Idealmente, un sistema de indicadores mide diversos componentes del sistema educativo, y ofrece también información sobre cómo interactúan los componentes singulares para producir el efecto de conjunto. En otras palabras, el conjunto de la información que ofrece un sistema de indicadores es más que la suma de sus partes. (Shavelson et al., 1989: 5-6)*

**PROPUESTA METODOLÓGICA PARA DESARROLLAR UN SISTEMA DE
INDICADORES EDUCATIVOS PARA EVALUAR LA CALIDAD DE LA
EDUCACIÓN EN MÉXICO**

Felipe Martínez Rizo

Seminario Internacional de Indicadores Educativos
Instituto Nacional para la Evaluación de la Educación
México, D. F., julio 9 de 2007

Introducción

En estas páginas se presenta una propuesta metodológica para el desarrollo de indicadores de la calidad del sistema educativo mexicano. La metodología es el fruto de un proceso comenzado tiempo atrás, con el desarrollo de un sistema de indicadores para la Universidad Autónoma de Aguascalientes, UAA (Cfr. Martínez Rizo et al., 1982). Después se propuso un sistema análogo para el conjunto del sistema educativo del estado de Aguascalientes (Martínez Rizo, 1996). Un esfuerzo más reciente, aprovechado en el trabajo del INEE, fue la tesis doctoral de Dora Ma. Ocampo (2004).

En esta versión se prescinde de muchas precisiones teóricas, que se pueden encontrar en un texto que presenta con mayor amplitud la propuesta metodológica en cuestión, al que se remite a las personas interesadas (Martínez Rizo, 2005). No se incluye tampoco una revisión histórica de los antecedentes de los indicadores sociales y educativos, para lo que se remite tanto al texto citado en la referencia anterior, como a uno de los anexos del proyecto de creación del INEE (Martínez Rizo, 2002)

Para entender la razón de esta propuesta, conviene recordar que la tarea evaluativa del INEE no se reduce a la elaboración y aplicación de pruebas del rendimiento escolar de los alumnos del sistema educativo mexicano, sino que debe atender otras dimensiones de la calidad educativa. El concepto de calidad que sustenta los trabajos del INEE así lo indica, como se verá más adelante.

De hecho, lo relativo a indicadores educativos constituye la primera de las tareas que se encomiendan al INEE en el Decreto Presidencial que le dio origen, publicado en el Diario Oficial de la Federación el 8 de agosto de 2002, que dice a la letra:

ARTÍCULO 3o.- Para el cumplimiento de sus objetivos, el Instituto colaborará con la Secretaría de Educación Pública en las evaluaciones que ésta deba realizar respecto del sistema educativo nacional, así como para la fijación de los lineamientos generales conforme a los cuales, las autoridades educativas locales deban evaluar sus respectivos sistemas educativos. Corresponderá al Instituto:

- I. Desarrollar y mantener en operación un sistema de indicadores que permita valorar en forma objetiva la calidad del sistema educativo nacional, en los niveles que le corresponden;

Las fracciones siguientes del artículo 3° del Decreto se refieren a las evaluaciones de los aprendizajes, de las escuelas, de programas, etc.

Sin embargo, la metodología para diseñar indicadores educativos no está bien sistematizada, en contraste con la de las pruebas de aprendizaje. En cuanto a éstas últimas, desde hace décadas existen programas educativos en los que los interesados en el campo de la psicometría pueden aprender todo lo necesario, y hay numerosas obras que proponen, con todo detalle y variantes menores, los pasos que deberán darse para generar buenos instrumentos. Un indicador claro de la madurez de las metodologías en lo relativo a pruebas es que se cuenta con estándares de calidad precisos al respecto, difundidos en textos específicos, publicados por las agrupaciones de profesionales del campo. (Cfr. AAHE, 1992; AERA, APA, NCME, 1985 y 1999; ETS, 2000; Martin, M. O. et al., 1999).

A partir de esos documentos de estándares de calidad para pruebas, hay incluso un trabajo mexicano en el mismo sentido, producido por el Centro Nacional de Evaluación para la Educación Superior, el CENEVAL. (Martínez Rizo et al., 2000)

No se encuentran, en cambio, textos similares para guiar el proceso de construcción de sistemas de indicadores. En algunos textos técnicos poco accesibles pueden encontrarse elementos (v. gr. Shavelson et al., 1991), pero no puede hablarse todavía de un cuerpo consolidado de orientaciones en este campo. Hay estándares de calidad relativos a estadísticas, educativas y otras, pero no específicamente en lo que se refiere a indicadores. (Cfr. Hofmann, L. M. 1991; OECD, 2004)

Lo anterior explica, entre otras cosas, que los sistemas de indicadores sean desarrollados, por lo general, en forma poco sistemática.

Aprovechando las experiencias mencionadas, el desarrollo del sistema de indicadores del INEE intentó seguir, desde 2003 y con éxito desigual, la metodología a la que se refiere este trabajo, en lo relativo a educación básica.

En 2005 y 2006, con la misma metodología se llevó a cabo la primera etapa del diseño de un *sistema nacional de indicadores*, que sí contempló el conjunto del sistema educativo y se hizo en colaboración con la Unidad de Planeación y Evaluación de Políticas Educativas de la Secretaría de Educación Pública. (Cfr. SEP-INEE 2006a y 2006b)

1. La noción de indicador

Aunque pueden rastrearse antecedentes en épocas antiguas, la construcción de *indicadores numéricos* sobre ciertos aspectos de la realidad social es un fenómeno relativamente reciente. En último término, este fenómeno se debe a la necesidad que tiene el hombre de utilizar *conceptos* más o menos precisos para captar el mundo que lo rodea, que inevitablemente deben limitarse a algunos de los aspectos o dimensiones de dicha realidad, ante la imposibilidad de conocerla intuitivamente en su totalidad; a esto se añade la ventaja de manejar datos numéricos y no sólo cualitativos.

El término indicador, sin embargo, es utilizado en sentidos no coincidentes, como puede verse con mayor amplitud en otra parte (Martínez Rizo, 2005).

Pese a ello hay coincidencias claras; la idea básica, la necesidad inevitable de recurrir a conceptos susceptibles de ser captados empíricamente, ante la dificultad de observar directamente dimensiones importantes y sutiles de la realidad, está presente desde los momentos iniciales de lo que luego se llamó *el movimiento de los indicadores*, como puede apreciarse en la cita siguiente, tomada de una obra que es considerada generalmente como la más representativa de esa época:

El problema clave de un sistema de indicadores sociales... es que nunca podemos medir directamente las variables que nos interesan, sino que tenemos que seleccionar substitutos en el lugar de esas variables. Podemos, por ejemplo, estar interesados en saber si una persona es o no ambiciosa. Pero no podemos observar la ambición en sí misma. Podemos hacer preguntas a la persona, y escuchar sus respuestas, o bien observar qué tan intensamente trabaja y en busca de qué tipo de retribución; a partir de ello podemos inferir si la persona es o no ambiciosa. (Bauer, 1966: 45)

Según Wyatt (1994) la definición más aceptada en la actualidad se deriva de la de Oakes, según la cual un indicador debe ofrecer, por lo menos, uno de los siguientes tipos de información:

- ◆ Sobre los logros de un sistema educativo en lo relativo a la obtención de determinados resultados; el indicador se liga, por tanto, a los objetivos, y constituye un punto de referencia para medir los avances (*benchmark*).
- ◆ Sobre algunas características que la investigación ha mostrado se relacionan con ciertos resultados; el indicador tiene así valor predictivo, porque su modificación traerá consigo otros cambios.
- ◆ Sobre rasgos centrales del sistema educativo (v.gr. insumos), o problemas, para entender cómo funciona.
- ◆ Sobre información relacionada con problemas, o sobre aspectos relevantes para la política educativa, que permitan apoyar la toma de decisiones. (Cfr. Oakes, 1986)

Siguiendo a Jaeger, que muestra la poca consistencia de las definiciones de indicadores que se encuentran en la literatura especializada (1978), Shavelson y colaboradores consideran que conviene adoptar una postura pragmática, y proponen la siguiente definición de trabajo: *un indicador es una estadística simple o compuesta que se relaciona con un constructo educativo básico y es útil en un marco de políticas públicas.* (1989: 4-5)

Al igual que Shavelson, Oakes distingue dos tipos básicos de indicador: los que consisten, respectivamente, en estadísticas simples o compuestas (1986: 3). En una clasificación más elaborada, el documento de la Comisión Europea ofrece una tipología de indicadores que los clasifica según seis dimensiones:

- ◆ Según el grado de procesamiento de la información de base que se requiera para construirlos, los indicadores pueden ser *elementales*, *derivados* (si relacionan dos indicadores elementales) o *compuestos* (si combinan varios indicadores elementales o derivados). Los indicadores pueden, además, *desagregarse* en varias categorías del mismo tipo (por ejemplo tasas de deserción por zonas), o *descomponerse* en categorías de tipo diferente, como por género.

- ◆ Según la comparabilidad de la información que contengan, podrán ser *específicos* (cuando sólo se usan en el marco de un programa particular, y por lo mismo no pueden compararse con los de otro programa), o *genéricos*, si se usan en varios programas, de manera que pueden hacerse comparaciones entre unos y otros.
- ◆ Según el alcance de la información que ofrezcan, puede haber indicadores de *programa*, si la información se refiere sólo a elementos internos, y de *contexto*, cuando se refiere a otros elementos.
- ◆ Según las fases del programa a que se refieran, podrán ser indicadores de *recursos* o insumos, de *salidas* o productos inmediatos, de *resultados* o productos mediatos, y de *impacto*, o productos de largo plazo.
- ◆ Según los criterios de evaluación o, si se prefiere, según las dimensiones del concepto de calidad, podrá haber indicadores de *relevancia*, *eficacia*, *eficiencia* o *desempeño*.
- ◆ Y según la cercanía del momento de obtención de la información con los procesos a evaluar, y el momento de su utilización, se podrá distinguir entre indicadores de *monitoreo*, para uso inmediato, o de *evaluación*, para usarse al final del proceso. (Cfr. Comisión Europea, 1999: 19-38)

La complejidad de los sistemas educativos y de los procesos que tienen lugar en su seno hacen que ningún indicador particular pueda ofrecer una imagen suficientemente amplia del conjunto, como para que sea útil para propósitos de evaluación. Por ello es importante distinguir entre indicadores singulares y *sistema de indicadores*. Como señalan diversos autores, un *sistema* no es la simple acumulación o yuxtaposición de indicadores particulares, sino un conjunto articulado de ellos, según cierta estructura. (Cfr. Oakes, 1986; Shavelson et al. 1987).

2. Metodología para el diseño de un sistema de indicadores

A partir de la experiencia propia y la literatura revisada (en especial Shavelson et al., 1989: 8-12; y Comisión Europea, 1999: 179-209), la metodología que el INEE propone utilizar para el diseño de sistemas de indicadores (Cfr. Martínez Rizo, 2005) comprende cuatro etapas básicas, como sigue.

Primera Etapa: elaboración del listado inicial de indicadores

El punto de partida es la idea, ya expresada, de que un sistema de indicadores no se reduce a un listado de ellos, simplemente yuxtapuestos. Para conformar un auténtico sistema, por el contrario, los indicadores deben integrarse lógicamente, según criterios precisos. Para identificar los componentes del sistema y los criterios de integración resulta útil la distinción que opone *objetos empíricos* y *objetos teóricos*.

Un sistema educativo comprende un gran conjunto de *objetos empíricos*: personas (aspirantes, alumnos, egresados, desertores, maestros, directores, entre otras); que se ubican en instituciones (planteles, agrupados en zonas, ubicadas en municipios y circunscripciones más amplias); en las que hay muebles, equipos y materiales (pupitres, libros de texto, cuadernos, pizarrones, computadoras y muchos más); que se usan en la enseñanza o la administración, con elementos menos tangibles: un currículo, cierta normatividad, algunas tradiciones, etc.

Cada uno de esos *objetos empíricos* constituye una realidad que, en sentido estricto, es inagotable: cada alumno tiene infinidad de rasgos personales, pertenece a una familia más o menos integrada, vive en cierta localidad, barrio o comunidad... Por consiguiente, los *objetos empíricos* mencionados como integrantes del sistema educativo, y el sinnúmero no mencionado, no pueden aprehenderse en su totalidad: antes de buscar información sobre ellos debemos precisar el propósito de nuestra indagación, para luego seleccionar los aspectos pertinentes para ese propósito. En otras palabras, debemos construir *objetos teóricos*, conformados por ciertas dimensiones, de especial interés para nuestros propósitos, de la inagotable realidad de los *objetos empíricos*.

Dado que el propósito que guía la construcción de un sistema de indicadores no es otro que el de evaluar la calidad del sistema educativo, precisar la noción de calidad es fundamental para que podamos pasar del *objeto empírico* al *teórico*. Dada la complejidad de los sistemas educativos, la noción de calidad aplicable no puede reducirse tampoco a una dimensión, sino que debe incluir al menos:

- ◆ **Relevancia**, entendida como la cualidad de un sistema educativo que define objetivos curriculares adecuados a las necesidades de la sociedad en que se sitúa.
- ◆ **Eficacia interna y externa**, entendiéndose por ello que el sistema educativo logre que una alta proporción de sus destinatarios acceda a la escuela, permanezca en ella hasta el final del trayecto previsto y egrese alcanzando los objetivos de aprendizaje.
- ◆ **Impacto**, que se dará cuando el sistema consiga que los aprendizajes alcanzados por los alumnos sean asimilados por ellos en forma duradera y se traduzcan en comportamientos fructíferos para la sociedad y los propios individuos.
- ◆ **Suficiencia y eficiencia**, si el sistema cuenta con recursos humanos y materiales suficientes y los aprovecha de la mejor manera, evitando despilfarros y derroches.
- ◆ **Equidad**, si tiene en cuenta la desigual situación de alumnos, familias, comunidades y escuelas, y ofrece apoyos especiales a quienes lo requieren, para que los objetivos educativos sean alcanzados por el mayor número posible.

Para relacionar elementos empíricos y dimensiones de la calidad educativa pueden utilizarse las nociones de la teoría de sistemas. De esta manera se podrá definir un verdadero *sistema de indicadores*, que considere insumos, procesos y productos, así como recursos y necesidades del entorno o contexto.

Una vez establecida la estructura del sistema podrá elaborarse un primer listado de indicadores, bien articulados entre sí. Se reitera la necesidad de evitar el error de diseñar un sistema de indicadores utilizando sólo información existente, aunque se diga que los aspectos no atendidos son también importantes. Esta salvedad pronto será olvidada y lo importante será definido por lo que cubran los indicadores disponibles. El sistema de indicadores deberá diseñarse a partir de lo que en principio se considera importante, aunque para muchos aspectos no se cuente con datos adecuados y deban realizarse después todos los esfuerzos para contar con la información necesaria.

Segunda Etapa: desarrollo de cada indicador del listado inicial

Cada indicador deberá desarrollarse luego, precisándose al menos los siguientes puntos:

- ◆ **Nombre**, lo más breve y claro que sea posible.
- ◆ **Definición**, con la mayor precisión y rigor conceptual.
- ◆ **Explicación**, si se requiere, la definición podrá desarrollarse, situando el indicador en un marco teórico pertinente. De especial importancia será precisar la ubicación de cada indicador en el conjunto del sistema, indicando la dimensión del concepto de calidad y los elementos sistémicos a los que corresponde: relevancia, eficacia, eficiencia, etc.; entorno, necesidades, insumos, procesos o productos.
- ◆ **Fuentes** de obtención de la información de base.
- ◆ **Fórmula de cálculo** y, en su caso, criterios para su estimación.
- ◆ **Desagregaciones** posibles (v. gr. por género, por entidad, por modalidad de escuela etc.)
- ◆ **Elementos para la interpretación** de los resultados.

Tercera Etapa: jueceo inicial de los indicadores desarrollados

Antes de organizar costosos procesos de recolección de datos para alimentar un sistema de indicadores, cada uno de los elementos que lo forman deberá someterse a escrutinio, mediante un procedimiento de jueceo entre expertos y usuarios, para valorar *a priori* el grado en que cada indicador parece satisfacer los criterios de calidad aplicables que son, al menos, los siguientes:

- ◆ **Validez**, como cualidad técnica básica. Sin llegar a distinciones finas sobre tipos o fuentes de validez, se entenderá de manera general como la propiedad del indicador que consiste en que mida realmente lo que se pretende que mida.
- ◆ **Confiabilidad**, la otra cualidad técnica esencial, consistente en que el indicador se defina conceptual y operacionalmente de forma tal que la información que permita obtener sea consistente a lo largo de las sucesivas aplicaciones del sistema.

- ◆ **Comparabilidad**, si el indicador es aplicable en contextos (educativos) diferentes en forma tal que permita comparaciones significativas.
- ◆ **Estabilidad temporal**, de modo similar, si el indicador permite comparaciones significativas a lo largo del tiempo.
- ◆ **Actualidad** de la información que dé el indicador (*freshness*), pues uno que no aporte información razonablemente reciente no es muy útil.
- ◆ **Sensibilidad**, que se dará si valores distintos de un indicador se asocian en forma consistente a diferencias significativas de los sistemas educativos.
- ◆ **Factibilidad de implementación**, en el sentido de que la información necesaria para construir un indicador pueda obtenerse. Cuando esa información no esté disponible deberá valorarse la importancia del indicador en relación con el costo que implicará.
- ◆ **Importancia**, en el sentido de que el indicador se refiera a un elemento del sistema educativo que tenga peso significativo en relación con alguna dimensión de la calidad.
- ◆ **Utilidad**, como el grado en que el indicador se refiera a un aspecto en el que las decisiones que tomen los responsables puedan impactar de manera significativa.
- ◆ **Claridad**, en el sentido de que la manera en que presente la información deberá facilitar una adecuada interpretación por parte de los usuarios.

Además de los criterios anteriores, aplicables a cada indicador, deberá valorarse la calidad del conjunto del sistema como tal, juzgando si comprende de manera suficiente y equilibrada todos los aspectos importantes del objeto de estudio con un número razonable de elementos, ni excesivo ni escaso, esto es los criterios de **Cobertura, Balance y Parsimonia** del sistema de indicadores como tal.

El análisis inicial *a priori* del conjunto de indicadores diseñado deberá involucrar a un número razonable de jueces, de tres tipos diferentes: *expertos* en el sentido de investigadores o especialistas de formación y experiencia teórica o disciplinaria sobre los temas cubiertos por el sistema; *usuarios especializados*,

en el sentido de responsables del sistema educativo con experiencia de nivel alto y medio (autoridades federales y estatales, jefes de sector, supervisores e inspectores); y *usuarios de base*, como directores de escuela, maestros y padres de familia.

El proceso de jueceo deberá comprender al menos tres vueltas, en forma similar a la que se emplea en la técnica *delphi*. En cada vuelta se solicitará a los jueces que califiquen cada indicador según los criterios de calidad en una escala ordinal de cinco valores, de muy adecuado a muy inadecuado, que se codificarán de 5 a 1. En cada vuelta se solicitará a los jueces expresar su opinión sobre los indicadores, ofreciéndoles información adicional de la siguiente manera:

- ◆ En la primera vuelta se solicitará la opinión de los participantes sobre cada uno de los aspectos mencionados de cada indicador (validez, confiabilidad, comparabilidad, etc.) ofreciendo la información de cada uno que resultó de la Segunda Etapa: nombre, definición, explicación, fuentes, fórmula de cálculo, desagregaciones posibles y elementos para la interpretación de los resultados.
- ◆ En la segunda vuelta se pedirá a los participantes que vuelvan a opinar sobre cada aspecto de cada indicador, teniendo a la vista información sobre las opiniones de los demás participantes en la primera vuelta, en la forma de medias y desviaciones de las puntuaciones obtenidas por cada indicador en cada criterio. Se les pedirá a quienes tengan opiniones que difieran mucho de la tendencia promedio que expresen argumentos para defender su posición.
- ◆ En la tercera se pedirá nuevamente la opinión de los participantes, pero dándoles además información sobre los argumentos de los participantes con opiniones extremas o discordantes en la vuelta anterior.

Con base en las opiniones de los jueces se podrá decidir cuáles indicadores satisfacen de manera más o menos amplia los criterios de calidad anteriormente mencionados. Entre una vuelta y otra podrán modificarse o descartarse aquellos indicadores que, según el consenso recabado, no satisfagan los criterios de calidad.

Convendrá dar más peso a las opiniones de algunos de los jueces en relación con ciertos criterios. Así, la opinión de los *expertos* deberá tener mayor peso para las decisiones sobre los criterios de validez, confiabilidad, comparabilidad y estabilidad temporal, así como para actualidad y sensibilidad; la opinión de los *usuarios especializados* es de especial relevancia para decidir sobre la factibilidad, importancia y utilidad; y la de los *usuarios de base* para juzgar sobre la claridad de los indicadores.

Teniendo en cuenta que algunos participantes podrán abandonar el proceso antes de terminar, y el tiempo que implica recoger las opiniones de cada vuelta y procesarlas antes de la siguiente, habrá que considerar un tiempo razonable para cada una (de dos a cuatro semanas) e incluir inicialmente a un número de jueces mayor al que se espera que termine el proceso, digamos un mínimo de 10 personas de cada una de las tres clases de participante.

Cuarta Etapa: prueba piloto o primera aplicación del sistema

Tras las etapas anteriores podrá procederse a una primera aplicación del sistema, alimentándolo con la información necesaria, sea previamente existente o bien generada *ex profeso*, y realizando los procesos de depuración y cálculo necesarios para obtener los primeros valores de cada indicador.

Con los resultados de la primera aplicación se podrá hacer un nuevo análisis de la calidad de cada indicador y del sistema en conjunto, esta vez *a posteriori*, revisando la consistencia de la información y contrastándola con otra comparable, por ejemplo derivada de investigaciones independientes. Podrá también valorarse en qué medida la información que arroja el sistema es interpretada correctamente y utilizada para sustentar acciones de mejora.

3. Pasos adicionales y maduración del sistema de indicadores

Tras las etapas anteriores el sistema estará en condiciones de operar, pero para su plena maduración deben considerarse dos pasos adicionales:

- ◆ El establecimiento y ajuste de referentes para contrastar los valores de los indicadores y llegar a juicios de valor.
- ◆ El refinamiento permanente de los indicadores que lo requieran, con base en la experiencia que se irá acumulando.

El establecimiento de referentes

Evaluar implica medir, pero no se reduce a ello. Un indicador captura o mide cierto aspecto de la realidad, lo que es indispensable para evaluarlo, pero no es suficiente: para llegar a un juicio sobre lo adecuado o inadecuado de lo medido hace falta un referente que defina lo que se considera adecuado o inadecuado.

Ese referente con que debe contrastarse el resultado de la medición es el *parámetro* o *estándar*. Es normativo, no empírico. Para definirlo, la pregunta que hay que formular es ¿con qué o con quién debemos compararnos? Para responder la pregunta hay que tener en cuenta que los referentes pueden ser de tres tipos: *óptimos*, *promedio* y *mínimos*.

Tratándose de juicios de valor sobre sistemas educativos nacionales, los referentes óptimos pueden ser los resultados educativos de los países más avanzados. Al interior de un sistema educativo, cuando se trate de juicios de valor sobre escuelas, los referentes óptimos serán los resultados de las mejores. Usar como punto de comparación tal tipo de referentes óptimos equivale al enfoque conocido como *benchmarking*, o la identificación de las llamadas *mejores prácticas*.

El uso de referentes promedio implica comparar los resultados de un sistema, una escuela o, en su caso, un alumno, con la media de los sistemas, escuelas o alumnos que se evalúa. Es el tipo de comparación que se refleja en expresiones como: *la escuela x está por encima de la media nacional*; o bien: *la media de México es inferior a la de los países de la OCDE*. Cuando, en las evaluaciones internacionales, se compara a un país como el nuestro con otros *de desarrollo similar*, de alguna manera se utiliza este tipo de referente.

El uso de referentes mínimos implica que se establezca el menor valor aceptable, cuyo cumplimiento permite, por ejemplo, que un alumno sea aprobado, o que una institución sea acreditada. El uso de los países de menor desarrollo como puntos de comparación podría verse como una aproximación al uso de referentes mínimos.

Cada referente (óptimo, promedio y mínimo) tiene cierto sentido e ilumina la realidad evaluada desde su perspectiva; ninguno es suficiente. Es conveniente usar los tres tipos de parámetro, para una valoración mas completa.

En el caso de las comparaciones internacionales, tiene sentido comparar la situación educativa de México con la de países más avanzados. La situación de estos últimos es diferente de la del nuestro, pero no deja de ser un punto de referencia, una meta a perseguir, aunque sea a mediano o largo plazo. La globalización del mundo contemporáneo lo hace inevitable. Utilizar sólo tales referentes, sin embargo, puede ser injusto, al no tener en cuenta numerosos factores demográficos, económicos, sociales y culturales. La equidad implica usar parámetros que consideren los contextos de las realidades evaluadas, comparando lo realmente comparable. Por ello el emplear también como referentes a los países de desarrollo similar, usando referentes *promedio*, es un complemento deseable en las comparaciones internacionales.

La evaluación de los sistemas educativos de las entidades federativas de nuestro país puede hacerse, de manera similar, en relación con referentes óptimos, promedio o mínimos. Como referente óptimo puede usarse el nivel de las entidades más desarrolladas, que se pueden considerar referentes de *benchmarking* para las menos avanzadas. La media nacional es el referente promedio natural. En cuanto a los referentes mínimos, las autoridades educativas podrán establecerlos. A nivel internacional, organismos como la UNESCO y la OCDE pueden fijar este tipo de parámetros mínimos.

Una forma diferente, y sugestiva, de definir referentes, es la de emplear como tales la situación del sistema a evaluar en el pasado, en el presente o en el futuro. La situación en cierto momento del pasado puede verse como referente mínimo. El futuro puede usarse como punto de referencia, en el sentido de que un sistema puede valorar si alcanza o no, si se aproxima o se aleja, a mayor o menor velocidad, a las metas que se haya establecido en determinado horizonte temporal. Posiblemente estos parámetros sean los más pertinentes para valorar la calidad educativa: una escuela o sistema escolar de calidad es, finalmente, aquella o aquel que siempre mejora respecto a sí mismo, sin idealizar el pasado y con metas ambiciosas pero realistas para el futuro.

Refinamiento de los indicadores y maduración del sistema

Es importante tener en cuenta que el desarrollo de indicadores de buena calidad implica necesariamente tiempos considerables.

Para asegurar la calidad los pasos metodológicos deben seguirse con cuidado, involucrando a expertos, pero no puede esperarse que los indicadores satisfagan plenamente los criterios de calidad desde la primera ocasión que se utilicen, sobre todo en los casos de los aspectos más complejos de la realidad a evaluar. Por ello la calidad implicará que los indicadores se revisen y corrijan una y otra vez, antes de que se les considere bien establecidos.

Si se tiene en cuenta el ritmo anual de los procesos de recolección de muchas de las estadísticas educativas, que se obtienen de las escuelas al inicio y/o fin de cada ciclo escolar, se comprenderá que el tiempo que implica no sólo el primer planteamiento de ciertos indicadores, sino su refinamiento hasta que alcancen un grado aceptable de madurez, implica necesariamente varios años.

Una de las especialistas más reconocidas en el tema, al sacar las lecciones del *movimiento de los indicadores sociales* de los años 60, señalaba:

Este análisis del diseño y la aplicación de indicadores... pueden darnos ideas sobre los errores que debemos evitar y las oportunidades que podemos aprovechar... Se necesita mucho tiempo para el desarrollo de un concepto satisfactorio, luego para la producción de una medida práctica y, finalmente, para asegurar que el indicador resultante sea entendido y utilizado. Los indicadores existentes han requerido 20, 30 o más años, desde el momento en que aparecieron las primeras demandas públicas y los primeros análisis teóricos...

Hay que contar con que la creación de un indicador nuevo toma largo tiempo. Este tiempo difícilmente puede abreviarse, dada la naturaleza iterativa de los procesos de formación de conceptos y de los esfuerzos de medición, así como la dificultad de identificar el carácter preciso de las preocupaciones públicas antes de producir una medida. Si se debe desarrollar un método nuevo es necesario además tener en cuenta el tiempo necesario para los ajustes por ensayo y error. Aun si ya existen algunos antecedentes de investigación y cierto consenso alrededor de un problema, es probable que pasen al menos 10 años entre el deseo de contar con una medida y la producción efectiva de esta. (de Neufville, 1975: 241-244)

Pensando en términos de *sistema de indicadores*, y no uno u otro aislado, no debemos esperar a tener perfectamente probado el primero para emprender el desarrollo del siguiente. Los tiempos de que habla de Neufville no deberán multiplicarse por el número de indicadores que queramos desarrollar, pero tampoco podrán acortarse. Por ello hay que tener en cuenta que el desarrollo de un sistema de indicadores maduro es una tarea de, al menos, una década.

4. El control de la calidad de los indicadores

Desde su nacimiento, el INEE ha tratado de que su trabajo de evaluación, tanto en lo relativo a pruebas de rendimiento como en los otros aspectos, constituya un avance cualitativo con respecto a lo hecho previamente en México en este terreno. Esta mejora debe incluir la articulación de las evaluaciones que llevan a cabo diferentes instancias y la difusión de sus resultados, frente a la situación de falta de coordinación y de opacidad prevaeciente hasta hace poco tiempo.

Pero la mejora fundamental, que condiciona las anteriores, es la que se refiere a la calidad técnica de las evaluaciones. El avance en este sentido es básico, porque la calidad de una evaluación es un requisito para que pueda articularse con otras y para que tenga sentido difundir sus resultados: si estos no son sólidos no podrán servir para sustentar decisiones de política ni para retroalimentar a las escuelas y los maestros.

Ahora bien: la calidad de una evaluación no resulta simplemente de un buen propósito; hay que poner los medios necesarios para conseguirla. La base para ello es el utilizar una metodología consistente, como la que se ha presentado en los apartados anteriores de este documento; es necesario además que se cuente con mecanismos apropiados de control de calidad.

Desde 2004, como parte de su primer Plan Maestro de Desarrollo, el INEE adoptó un *modelo de calidad* (INEE, 2004) que comprende tres componentes:

- ◆ **Criterios de calidad precisos**, a cumplir en cada paso de una evaluación.
- ◆ **Procedimientos claros** a seguir para asegurar se cumplan los criterios.
- ◆ **Instancias internas y externas**, integradas por personas calificadas de un perfil adecuado, que supervisen la operación de los procedimientos y el cumplimiento de los criterios.

Para establecer su modelo el INEE no partió de cero, sino que se basó en la experiencia de las mejores prácticas internacionales en el campo, que se plasman en documentos como los citados antes. (AAHE, 1992; AERA, APA, NCME, 1985 y 1999; ETS, 2000; Martin, M. O. et al., 1999).

Se partió, pues, de la convicción de que lograr buena calidad técnica no puede resultar sólo de la buena voluntad; había también conciencia de que tampoco es el producto mágico de ingredientes misteriosos que no pueden definirse: la calidad sólo puede ser el fruto de un trabajo intenso y disciplinado, que se apegue a **criterios** decantados por la experiencia, mediante **procedimientos** explícitos, y con la supervisión de **instancias de control** apropiadas para prevenir desviaciones y, en su caso, detectarlas y corregirlas. Para hacer realidad lo anterior se necesitan dos tipos de instrumentos:

- ◆ **Manuales Técnicos** que describan los criterios de calidad a respetar y los procedimientos a seguir, así como los controles que vigilarán que los procedimientos se sigan y los criterios de calidad se respeten.
- ◆ **Reportes Técnicos** que den cuenta de la manera y medida en que se cumplieron los criterios y se siguieron los procedimientos, de manera que las instancias internas y externas de control puedan verificarlo.

No parece superfluo añadir que el INEE considera que este modelo, basado en las mejores prácticas del campo, es más apropiado para asegurar la calidad de las evaluaciones que un modelo general como el que se conoce con las siglas ISO, en sus diferentes versiones. Si bien el espíritu de ambos modelos es similar, los intentos de aplicar un modelo desarrollado en el medio industrial a campos de alta especialidad como el de la evaluación no dan resultados satisfactorios, e ignoran los importantes avances que las comunidades de los especialistas han conseguido a lo largo de décadas.

Los criterios de calidad se aplican de manera distinta a los diversos tipos de evaluaciones, como pruebas, indicadores o evaluaciones de procesos. Por ello el modelo de calidad del INEE comprende criterios generales, aplicables en principio a toda evaluación, y criterios particulares para los tipos que maneja el Instituto. El cuadro siguiente, tomado del documento que contiene el modelo completo (INEE 2004), sintetiza los criterios aplicables a los indicadores.

ESTÁNDARES DE CALIDAD PARA SISTEMAS DE INDICADORES

ASPECTOS	PRODUCTOS	CRITERIOS	PROCEDIMIENTOS
Planeación-conceptualización definición de dimensiones	Documento con definición del sistema	Cobertura, Balance Parsimonia	Análisis documental por especialistas
Diseño de indicadores individuales	Manual Técnico. Docto. c/indicadores iniciales (fichas técnicas)	Validez, Confiabilidad Comparabilidad Sensibilidad, Factibilidad Importancia	Validación inicial <i>a priori</i> mediante jueceo con especialistas y usuarios
Jueces inicial de indicadores	Resultado de jueceo Indicadores ajustados (fichas modificadas)	Utilidad	
Prueba piloto de indicadores o aplicación inicial	Documento con análisis <i>a posteriori</i> de datos; Indicadores ajustados	Los anteriores más: Actualidad, estabilidad temporal y claridad	Validaciones <i>a posteriori</i> Análisis de series Juecos periódicos
Obtención, certificación y generación de información	Manual Técnico Reporte Técnico	Calidad instrumentos y procesos; Confiabilidad de resultados	Monitoreo de procesos Análisis de Reporte T. Estudios <i>ad hoc</i>
Integración de información, creación y mantenimiento de bases de datos	Manual Técnico Bases de datos	Apego a Manual Nº y tipo errores de bases	Monitoreo de procesos Controles <i>ad hoc</i> Análisis reporte y bases
Construcción de series datos	Series construidas	Consistencia resultados	Análisis documentales
Producción de informe anual	Informes terminados	Cobertura contenido Precisión análisis Objetividad juicios Claridad de expresión	Análisis documentales
Difusión de información	Reportes publicados (en línea-papel) Materiales de divulgación para diversas audiencias	Acceso a información: autoridades, maestros y directores, padres de familia, investigadores. Uso de resultados	Análisis documental
Desarrollo, actualización y mantenimiento del sistema	Docto. con definiciones actualizadas	Cobertura, balance y parsimonia del sistema	Análisis documental por especialistas

En su versión completa, el cuadro incluye las instancias responsables del control de la calidad de los indicadores. La instancia interna de control no es otra que el Consejo Técnico del Instituto; las instancias externas serán los grupos de especialistas de cada uno de los aspectos de que se trate.

Conclusión

La obra antes citada sobre indicadores de la Comisión Europea tiene una reflexión sobre las lecciones de la experiencia internacional:

La experiencia internacional en el campo de los indicadores es amplia y ha existido por largo tiempo. En el origen de los desarrollos en este dominio encontramos el éxito de unos indicadores macroeconómicos importantes: el producto interno bruto, el índice de precios o la tasa de desempleo. En los años 70 este éxito hizo surgir el desarrollo de indicadores sociales, actividad que implicó considerable energía y fue también la fuente de un malestar igualmente grande. (1999: 219)

Sheldon afirma que *con cada movimiento de indicadores, el entusiasmo creció y disminuyó, a medida que el optimismo sobre lo que podrían conseguir los indicadores dejó el lugar a la realidad de lo que efectivamente consiguieron, tanto en la educación como, más generalmente, en la sociedad.* (Según Shavelson et al., 1989: 2)

Como muestran esos especialistas, tras varias experiencias frustrantes las expectativas sobre lo que pueden aportar o no los sistemas de indicadores son más realistas. La literatura muestra que se ha llegado a un consenso en cuanto a que los indicadores no pueden, por sí mismos, fijar objetivos o prioridades, evaluar programas o establecer balances. Todo ello puede apoyarse en la información de un buen sistema de indicadores, pero implica otras acciones. Las expectativas sobre los indicadores que hay ahora en los sistemas educativos con mayor experiencia al respecto son bastante modestas: *pueden servir para describir y plantear problemas con mayor precisión; para detectarlos más tempranamente; para tener pistas sobre programas educativos prometedores y cosas similares.* (Shavelson et al., 1989: 7-8)

La obra citada de Judith Innes de Neufville (1975) tuvo reimpressiones en 1989 y 1994, y una “segunda edición expandida” en 2002. La persistencia del interés en su contenido a lo largo de casi tres décadas muestra su solidez, y las consideraciones de la autora en la segunda edición permiten apreciar el cambio de perspectiva que se produjo entre los estudiosos del tema a lo largo del tiempo. La autora señala que casi todos los aspectos técnicos de la versión de 1975 de su libro siguen vigentes, lo que justifica la reimpresión del texto original. Al mismo tiempo, la investigadora considera que lo técnico se enmarca ahora en una visión de la relación entre conocimiento y toma de decisiones mucha más rica que la que prevalecía en la década de 1970, todavía marcada por una concepción ingenuamente positivista. (Innes de Neufville, 2002)

Esta se caracterizaba por concebir una relación simple, lineal. En ella los indicadores eran el fruto de un trabajo técnico puro, en el que no interferían los intereses de grupo ni consideraciones extraacadémicas. Sobre la base del conocimiento sólido de la realidad que tales indicadores proporcionaban, los políticos podían fácilmente tomar decisiones plenamente racionales; si no lo hacían, era debido a la corrupción o la mezquindad.

Dos décadas más tarde, la autora enmarca sus concepciones técnicas sobre indicadores en una visión más compleja de la relación entre conocimiento y acción, a partir de los resultados de sus investigaciones y con apoyo en posturas como el viejo pragmatismo de principios del siglo XX, las teorías interpretativas, las concepciones de Berger y Luckmann y las de Habermas.

En esta perspectiva el conocimiento útil no surge simplemente del trabajo técnico: se mezcla inexorablemente con puntos de vista que privilegian unos aspectos de la realidad y se confrontan o complementan de manera compleja. La utilidad de un sistema de indicadores para apoyar decisiones no depende únicamente de su solidez técnica, que no debe descuidarse, sino también de la legitimidad que le dé la participación de actores varios en su construcción y de la riqueza del proceso mismo, con aportes técnicos, discusiones que aclaren puntos difíciles y consensos más o menos amplios laboriosamente alcanzados.

En el campo educativo, un sistema de indicadores no debe ser visto como una panacea que remediará rápidamente los males de los sistemas escolares. Lo que puede ofrecer es menos sensacional, pero no despreciable, sobre todo si se tienen en cuenta las limitaciones de las estadísticas existentes. Sin olvidar la complejidad de la construcción de consensos, un buen sistema de información educativa es necesario, aunque no suficiente, para que las decisiones tengan base sólida; pese a ello pocos países cuentan con uno adecuado, sobre todo en el llamado tercer mundo, en parte, sin duda, por la dificultad técnica para desarrollarlo en forma paralela al enorme crecimiento que experimentaron los sistemas escolares en todo el mundo a lo largo de la segunda mitad del siglo XX. Hace poco más de una década Puryear decía al respecto:

...Las estadísticas sobre educación de aproximadamente la mitad de los estados miembros de la UNESCO, incluyendo al menos cinco de los nueve países en vías de desarrollo más grandes, tienen lagunas y debilidades serias. Con mucha frecuencia las bases de datos sobre el tema carecen de confiabilidad. Aún en los países industriales avanzados los datos sobre costos y gastos educativos son muy deficientes. En casi todos los países los datos sobre analfabetismo no son confiables. Y los expertos opinan que las estadísticas educativas de 20 o 30 países son un verdadero desastre. (1993: 4)

La mayoría de los países carece de medidas sistemáticas sobre resultados educativos y las evaluaciones comparativas internacionales son raras. Puryear señalaba que suele haber estadísticas mejores sobre cuestiones económicas, demográficas o de salud, y precisaba:

*...Las estadísticas educativas existentes tienen un enfoque muy estrecho. Dejan fuera las medidas de calidad, de procesos y de productos. Los gobiernos han centrado la atención en un sólo enfoque, bastante estrecho, para monitorear sus sistemas educativos -los conteos de tipo censo- y en un sólo tipo de datos: los relativos a insumos. Se han concentrado casi enteramente en registrar el número de profesores, alumnos y edificios en el sistema formal, y en la importancia del gasto público destinado a la educación. Casi no han prestado atención a documentar cómo funcionan las escuelas o qué aprenden los alumnos... Un buen indicador de la existencia de estadísticas educativas de muy buena calidad, lo constituye la capacidad de reportar datos sobre la edad de los alumnos, que son indispensables para calcular tasas **netas** de matrícula. Únicamente unos 60 de los 175 estados miembros de la UNESCO reportan tales datos en forma consistente. (1993: 6)*

Poco antes, los primeros pasos de la OCDE para desarrollar su sistema de indicadores educativos comenzaban con un diagnóstico semejante:

Aunque muchos países publican impresionantes anuarios estadísticos, la cantidad y la calidad (validez, consistencia, etc.) de los datos recolectados es sumamente desigual, no sólo de un país a otro, sino incluso dentro de un mismo país. La Conferencia de la OCDE celebrada en Washington en 1987 señaló que, aunque la mayor parte de los 24 países participantes reportó que recababa datos para construir algunos indicadores educativos, su tipo y uso variaba mucho, y que muy pocos países estaban en condiciones de ofrecer conjuntos completos. Además, muchas presentaciones coincidieron en señalar problemas similares: datos incompletos o faltantes; falta de confiabilidad, debida en parte a técnicas de muestreo deficientes; insuficiente validez de las variables de contexto; dificultad para decidir qué datos recoger; problemas en cuanto al control de la información y el acceso a ella.

...Aún los sistemas más completos de estadísticas nacionales presentan hoyos negros y lagunas de información en temas de los que casi no hay datos, como costos unitarios, gasto educativo privado o la contribución económica de empresas y familias al gasto educativo total. La mayoría de los sistemas son notablemente débiles en cuanto a información sobre el conocimiento que tienen los maestros de las materias que deben enseñar, el aprendizaje de los alumnos, en especial en niveles cognitivos altos, o los procesos que tienen lugar dentro de la escuela. (CERI-INES 1991: 8-9)

Por la misma época, el grupo de expertos encargado por el gobierno de los Estados Unidos del diseño de un sistema de indicadores educativos a escala nacional identificaba problemas semejantes, que muestran que la pobreza y la inconsistencia de las estadísticas educativas no son exclusivas de los países menos desarrollados:

Algunos problemas técnicos serios de los indicadores actualmente existentes, junto con grandes lagunas en las fuentes de datos disponibles, plantean problemas formidables a la tarea de construir un sistema de indicadores educativos en el nivel nacional. El Council of Chief State School Officers, a solicitud del National Center for Education Statistics, ha analizado las fallas de medidas de resultados educativos comparables, observando por ejemplo que hasta 1987 los estados americanos empleaban al menos diez maneras distintas de contar sus escuelas y unos doce métodos diferentes de reportar las cifras de matrícula. Los datos sobre deserción son notoriamente poco confiables y representan unos de los datos más ambiguos que se reportan sobre la educación norteamericana. (Special Study Panel, 1991: 15)

Este texto, y en especial su conclusión, no buscan desalentar a quienes emprendan la tarea de desarrollar un sistema de indicadores. Lo que pretenden es apoyar sus esfuerzos, de dos maneras: proponiendo una metodología sencilla y clara para tal propósito, y ayudando a que tanto los técnicos como los tomadores de decisiones y los usuarios de la información tomen conciencia de la complejidad de la tarea, de manera que permitan que transcurran los plazos necesarios para que se haga bien.

Para terminar en un tono más optimista concluyamos que, para alcanzar una meta ambiciosa, hay que aceptar las limitaciones del punto de partida y mejorar progresivamente; y que para ello conviene institucionalizar mecanismos de corrección permanente, indispensables para avanzar en la dirección correcta, como apuntan las dos últimas citas:

...se presta considerable atención a la imprecisión de muchos datos estadísticos, a las distorsiones incorporadas sistemáticamente en algunos casos, a los frecuentes errores de interpretación y a la creciente posibilidad de manipular los datos... para tener mejor información sobre la calidad de vida se necesitan mejoras tanto en la cantidad como en la calidad de las estadísticas... La conclusión parece ser que en vez de no hacer nada, es preferible comenzar con datos malos, advertir a todos sobre sus defectos y limitaciones, y buscar una mejora gradual gracias a su uso. (Gross, 1966: xvi)

Propongo que se establezca una Comisión de Estadísticas, que tenga la responsabilidad de auditar las series de datos, con capacidad para certificar los productos estadísticos... al auditar series de datos importantes, la Comisión debería hacer muestreos para verificar la precisión de las bases de datos y de los cálculos, y debería examinar cuidadosamente los procedimientos establecidos para ver si se respetan escrupulosamente. (Wickens, 1953; citado por Biderman 1966: 129-130)

BIBLIOGRAFÍA.

- AAHE (1992). "Principles of Good Practice for Assessing Student Learning". *The AAHE Assessment Forum*. Washington. American Association for Higher Education.
- AERA, APA, NCME (1999). *Standards for educational and psychological testing*. Washington. American Educational Research Association.
- AERA, APA, NCME (1985). *Standards for educational and psychological testing*. Washington. American Psychological Association.
- BAUER, RAYMOND A. Ed. (1966). *Social Indicators*. Cambridge, MIT.
- BIDERMAN, ALBERT D. (1966). Social Indicators and Goals. En BAUER, RAYMOND A. Ed., 1966: 68-153.
- CERI-INES (1991) *An International Handbook of Educational Indicators*. Paris. OCDE.
- COMISION EUROPEA (1999). *Selection and use of indicators for monitoring and evaluation. Evaluating socio-economic programmes, Vol. 2*. Luxembourg. Office for Publications of the European Communities.
- ETS (2000). *ETS Standards for Quality and Fairness*. Princeton, N. J. Educational Testing Service.
- GROSS, BERTRAM M. (1966). Preface. A Historical Note on Social Indicators. En BAUER, RAYMOND A. Ed., 1966: ix-xviii.
- HOFFMAN, LEE M. (1991). *Standards for Education Data Collection and Reporting*. Washington. National Center for Education Statistics.
- INEE (2004). "El modelo de calidad del INEE. Estándares técnicos para las evaluaciones educativas". *Hacia un nuevo Paradigma para la Evaluación Educativa. La Perspectiva del INEE*. México. INEE, pp. 11-24.
- INNES DE NEUFVILLE, JUDITH (2002). *Knowledge and Public Policy: The Search for Meaningful Indicators*. New Brunswick. Transaction Publishers. 2a ed. Expandida de *Social Indicators and Public Policy*.
- (1975). *Social Indicators and Public Policy. Interactive Processes of Design and Application*. Amsterdam. Elsevier Scientific Publishing Co.

- JAEGAER, RICHARD M. (1978). About educational indicators: Statistics on the conditions and trends in education. *Review of Research in Education*. Vol. 6, 276-315.
- MARTIN, MICHAEL O. et al. (1999). *Technical Standards for IEA Studies*. Amsterdam. International Association for the Evaluation of Educational Achievement.
- MARTÍNEZ RIZO, FELIPE (2005). El diseño de sistemas de indicadores educativos: consideraciones teórico-metodológicas. *Cuadernos de Investigación*, N° 14. México, INEE.
- (2002). Revisión de literatura sobre evaluación de sistemas educativos. En VV, *La calidad de la educación en México: Perspectivas, análisis y evaluación*. México. M.Á. Porrúa-SEP-Congreso de la Unión, p. 305-427.
- (1996). *La calidad de la educación en Aguascalientes. Diseño de un sistema de monitoreo*. Aguascalientes. UAA-IEA.
- et al. (2000). *Estándares de Calidad para Instrumentos de Evaluación Educativa*. México. CENEVAL.
- et al. (1982). *La planeación universitaria. Una metodología para universidades mexicanas*. México, UNAM.
- OAKES, JEANNIE (1986). *Educational Indicators: A Guide for Policymakers*. New Brunswick Center for Policy Research in Education, Rutgers University; The Rand Corporation; University of Wisconsin-Madison.
- OCAMPO, DORA Ma. (2004). *Un sistema de indicadores para instituciones mexicanas de educación superior*. Tesis doctoral no publicada, dirigida por Felipe Martínez Rizo. Universidad Autónoma de Aguascalientes.
- OECD (2004). *OECD Handbook of Internationally Comparative Education Statistics: Concepts, Standards, Definitions and Classifications*. Paris. Organization for Economic Cooperation and Development.
- PURYEAR J.M. (1993) "Status and Problems of International Educational Statistics and Research". CIES Newsletter, No. 103:4.

- SEP-INEE (2006a). *Sistema de Indicadores Educativos de los Estados Unidos Mexicanos. Conjunto básico para el ciclo escolar 2004-2005*. México. SEP-INEE.
- (2006b) *Plan de Desarrollo del Sistema Nacional de Indicadores Educativos*. México. INEE-SEP. Puede consultarse en:
http://www.inee.edu.mx/images/stories/Sistemaindicadores_pdf/plan%20de%20desarrollo%20documento.pdf
- SHAVELSON, RICHARD J., LORRAINE M. McDONNELL y JEANNIE OAKES (1991). Steps in designing an indicator system. *Practical Assessment, Research & Evaluation*. Electronic Journal. ERIC Clearinghouse on Assessment and Evaluation.
- SHAVELSON, RICHARD J.; LORRAINE McDONNELL; JEANNIE OAKES Eds. (1989). *Indicators for Monitoring Mathematics and Science Education. A Sourcebook*. Santa Monica, Ca. Rand Corporation.
- SHAVELSON, RICHARD J.; LORRAINE McDONNELL; JEANNIE OAKES; NEIL CAREY; con LARRY PICUS (1987). *Indicator Systems for Monitoring Mathematics and Science Education*. Santa Monica, Ca. Rand Corporation.
- SPECIAL STUDY PANEL ON EDUCATION INDICATORS. (1991) *Education Counts. An Indicator System to Monitor the Nation's Educational Health*. Washington. National Center for Educational Statistics.
- TUIJNMAN A. y N. BOTTANI (Eds.) (1994) *Making Education Count. Developing and Using International Indicators*. Paris. OCDE.
- WICKENS, A. J. (1953). Statistics and the Public Interest. *The Journal of the American Statistical Association*. Vol. 48, pp. 1-14.
- WYATT T. (1994). Education indicators: a review of the literature. En TUIJNMAN y BOTTANI, 1994:99-121.