

El Diseño de Sistemas de Indicadores Educativos: Consideraciones Teórico-Metodológicas

Felipe Martínez Rizo

CUADERNO No. 14



Instituto Nacional para la Evaluación de la Educación

**COLECCIÓN CUADERNOS
DE INVESTIGACIÓN**

ISSN 1665-9457

El Diseño de Sistemas de Indicadores Educativos: Consideraciones Teórico-Metodológicas

Felipe Martínez Rizo*

CUADERNO No. 14

*Director General del Instituto Nacional para la Evaluación de la Educación

Este texto puede consultarse en: www.inee.edu.mx

MÉXICO, AGOSTO, 2005

CONTENIDO

♦ Introducción	3
♦ Antecedentes históricos	4
♦ Precisiones terminológicas y conceptuales	6
♦ Metodología para el diseño de un sistema de indicadores	11
-Primera etapa: elaboración del listado de indicadores	11
-Segunda etapa: desarrollo de cada indicador	12
-Tercera etapa: jueceo inicial	13
-Cuarta etapa: prueba piloto o primera aplicación	15
-Etapas posteriores	15
♦ La maduración de los sistemas de indicadores	17
♦ Conclusión	18
♦ Bibliografía	22

En una profunda afirmación que todavía resuena en los oídos de científicos y filósofos, Emmanuel Kant señaló en una ocasión que un percepto sin concepto es ciego, pero un concepto sin percepto es vacío.
(Gross, 1966: 255-256)

INTRODUCCIÓN

A diferencia de lo que ocurre con las pruebas de aprendizaje, la metodología para diseñar indicadores educativos no está suficientemente sistematizada o codificada.

En lo relativo a pruebas, desde hace medio siglo se dispone de manuales que explican con detalle los pasos a dar para elaborar instrumentos adecuados. Desde luego, no es sencillo seguir al pie de la letra todas las indicaciones técnicas, pero si se consigue hacerlo en un grado aceptable, el resultado será un instrumento razonablemente válido y confiable, aun si no se incorporan los desarrollos de las últimas décadas, como la teoría de la respuesta al ítem.

En cambio, no se encuentran textos similares para guiar el proceso de construcción de sistemas de indicadores. Los escasos apuntes existentes al respecto deben buscarse en documentos técnicos y artículos poco accesibles, pero no puede hablarse todavía de un cuerpo estándar de orientaciones en este campo. Lo anterior es uno de los factores para explicar que muchos de los sistemas de indicadores en nuestro medio, sobre tema educativo u otro, sean el resultado de procesos poco rigurosos de diseño, y adolezcan de deficiencias graves.

El trabajo del Instituto Nacional para la Evaluación de la Educación (INEE) en 2003 incluyó, con base en las funciones que le asigna su Decreto de Creación, el desarrollo de una primera versión de un sistema de indicadores educativos. Para llevar a cabo adecuadamente esta tarea, se siguió un procedimiento sistemático, buscando aprovechar al máximo la experiencia previa sobre el tema.

En las páginas siguientes se sistematiza la metodología utilizada, precedida por algunas precisiones teóricas. El texto retoma elementos de trabajos previos, en especial la revisión de literatura hecha como preparación del Proyecto del INEE (Martínez Rizo, 2002) y el trabajo de tesis que desarrolla Dora Ma. Ocampo, bajo la dirección del autor, en el Doctorado Interinstitucional en Educación de la Universidad Autónoma de Aguascalientes.

ANTECEDENTES HISTÓRICOS

La relación entre indicadores y estadísticas es muy clara al grado que, como veremos, los primeros pueden definirse como un tipo especial de las segundas. En ambos casos estamos ante esfuerzos por concretar nociones abstractas, o por cuantificar lo cualitativo.

El sentido original de la palabra estadística muestra que la idea de cuantificar fenómenos sociales es antigua. En el siglo XVII el término designaba información sobre algún aspecto de la situación del Estado, el cual podía consistir en datos cualitativos o numéricos; luego el sentido se reservó a la información cuantitativa.

Desde hace siglos se utilizan medidas burdas de ciertos aspectos de la vida pública, pero es sólo en épocas más recientes cuando la cuantificación se manifiesta de manera más amplia. Inicialmente ocurrió en relación con aspectos demográficos o económicos, los cuales se prestan de manera especial para su tratamiento en forma numérica. El desarrollo de la noción de *indicador social* surge después, como resultado de la necesidad de complementar las cifras de población o los datos económicos con información sobre otras dimensiones de la vida social.

En un estudio temprano sobre los indicadores, Bidernan (1966) muestra que, sin ser designados con ese término, estimaciones numéricas de ciertos fenómenos sociales fueron usadas desde hace mucho por los tomadores de decisiones, tanto para establecer metas precisas como para valorar su cumplimiento. Desde fines del siglo XVIII los presidentes de los Estados Unidos de América utilizaron ese tipo de estimaciones en sus informes anuales (Mensajes sobre el Estado de la Unión). El número de estas medidas (*indicadores*) pasó de un promedio de tres en cada informe de fines del siglo XVIII a 11 a mediados del siglo XX. Según el mismo autor, la tendencia se afirmó claramente a partir de 1933: desde entonces las expresiones oficiales de los objetivos del desarrollo nacional se acompañan regularmente de medidas numéricas.

En el campo educativo, el interés por el desarrollo de indicadores se ha manifestado en reiteradas ocasiones en la historia de los Estados Unidos, en relación con preocupaciones por la calidad de las escuelas. Según Shavelson y colaboradores, al terminar la Guerra de Secesión la preocupación se manifestó en el *Common School Movement*. Según estos autores, la creación del Departamento de Educación, en 1867, se justificó, entre otras razones, por la necesidad de elaborar indicadores (1989: 1).

Los estudiosos del tema coinciden en señalar la importancia que tuvo en ese sentido la iniciativa del Presidente Hoover, quien en 1929 formó una comisión de especialistas, con la tarea de desarrollar un estudio sobre las tendencias sociales en los Estados Unidos. El informe derivado del trabajo de esa comisión, publicado en 1933 con el título *Tendencias sociales recientes en los Estados Unidos* (Comité Presidencial, 1933), es la obra que inaugura la extensa bibliografía sobre indicadores sociales.

En Estados Unidos una nueva ola de interés sobre el tema se dio en la década de 1960, en la estela de la lucha por los derechos civiles, y en el marco de los programas de *la Gran Sociedad y Lucha contra la Pobreza* del Presidente Lyndon B. Johnson. La importancia de esta tendencia fue tal que para designarla se acuñó la expresión *movimiento de los indicadores sociales*. La abundancia de la literatura producida en esa época se pone de manifiesto en bibliografías temáticas como las de Wilcox *et al.* (1972) y Gilmartin *et al.* (1979), o en un manual como el de Russett, Alker, Deutsch y Lasswell (1965). El interés se extendió a varios países europeos y a organismos internacionales como la ONU y la OCDE. *Cfr.* obras como la de Delors, 1971; Shonfield y Shaw, 1972; CERI-INES, 1973; Departamento de Asuntos Económicos y Sociales de la ONU, 1975.

El *movimiento de los indicadores* de los años sesenta del siglo XX tuvo un componente educativo importante, a consecuencia del impacto que tuvo la puesta en órbita, por la Unión Soviética, del primer satélite artificial, lo cual se interpretó como síntoma del retraso de la enseñanza de matemáticas y ciencias en los Estados Unidos respecto a la URSS. En el plano internacional, los primeros esfuerzos por realizar evaluaciones comparables del aprendizaje alcanzado por alumnos de diversos países fue otra consecuencia del *Sputnik*, con los primeros trabajos de la *International Association for the Assessment of Educational Achievement* (IEA).

El *Special Study Panel on Education Indicators* observa que el famoso informe de la *National Comisión on Excellence in Education*, publicado en 1983 con el título de *Una nación en peligro*, se basó en comparaciones internacionales del desempeño de los alumnos de diversos países, y que un año después el Secretario de Educación comenzó la difusión de *The Wall Charts* donde publican anualmente algunos indicadores de los sistemas educativos estatales. (1991: 11).

En el mismo año de 1983, una comisión del *National Science Board* sobre la educación preuniversitaria de matemáticas y ciencias, recomendó al gobierno federal que financiara y mantuviera un mecanismo nacional para medir el desempeño de los alumnos, en una forma que permitiera evaluaciones nacionales, estatales y locales, así como comparaciones sobre el progreso educativo. (Shavelson *et al.*, 1989: 2).

A partir de la década de 1980 la preocupación por la calidad educativa alcanzó niveles sin precedentes en muchos países, ante la creciente competencia económica internacional. En ese marco adquieren sentido los esfuerzos por tener evaluaciones educativas más completas y confiables, lo cual incluye mejores sistemas de indicadores.

En Estados Unidos estas preocupaciones se reflejaron en diversos esfuerzos estatales a lo largo de la década de los ochenta (*cfr.* Odden, 1990); en las Metas 2000, adoptadas por el presidente tras la Cumbre Educativa de los gobernadores en octubre de 1989; en el establecimiento de estándares nacionales, el reforzamiento de los sistemas de pruebas y, finalmente, en las disposiciones de la ley *No Child Left Behind* firmada por el Presidente Bush

a principios de 2002. En todos los casos el desarrollo de mejores sistemas de indicadores formó parte destacada de los esfuerzos de evaluación. Referencias útiles para seguir el proceso en las últimas décadas son el estudio de Kaagan y Loby (1989) sobre los sistemas estatales de indicadores; los textos de Oakes (1986 y 1989); los de Shavelson y colaboradores (1987 y 1989); la propuesta del *Special Study Panel on Education Indicators* (1991); el artículo de Porter (1991); y el documento del Banco Mundial (Greaney y Kellaghan, 1996).

Muestras del interés reciente por los indicadores fuera de Estados Unidos son los textos ingleses de Fitz-Gibbon (1996) y Tymms (1999); así como los de Sauvageot sobre los sistemas de indicadores de dos países africanos (1992 y 1993). En el campo de las organizaciones internacionales, sobresale el trabajo de la OCDE, que en los años noventa retomó el proyecto de indicadores iniciado dos décadas antes y lo llevó a su actual etapa de madurez. *Cfr.* Bottani y Delfau, 1990; CERI-INES 1991a, 1991b, 1992, 1995a y 1995b; Tuijnman y Bottani, 1974. Más recientes, pero no menos importantes, son los esfuerzos de la Comisión Europea (1999).

El proyecto *Indicators of National Education Systems* (INES) comenzó a gestarse en noviembre de 1987 en una reunión cuyo propósito era discutir nuevos enfoques para desarrollar estadísticas educativas comparables. En 1988 la OCDE, a través del *Center for Educational Research and Innovation* (CERI), aprobó el proyecto que pronto adquirió grandes dimensiones: además de un grupo técnico integrado por medio centenar de expertos y por los coordinadores nacionales de los países miembros, se formaron cuatro redes de expertos para trabajar temas específicos: una sobre los resultados educativos de los estudiantes; otra sobre el destino de los estudiantes en el mercado laboral y la sociedad; una más sobre las características de las escuelas y los procesos que tienen lugar en su interior; y una cuarta en relación con las opiniones, expectativas y actitudes de la población sobre la educación. (Tuijnman y Bottani, 1994: 25).

El principal resultado del INES es la publicación *Education at a Glance*, que recoge anualmente los indicadores educativos de los países miembros de la OCDE. El desarrollo del sistema de indicadores educativos sigue su marcha: las ediciones anuales de *Education at a Glance* incluyen información cada vez más rica sobre contexto, costos, recursos y procesos escolares y resultados de la educación. Además el programa se extendió hacia países que no son miembros de la OCDE, con el programa *World Education Indicators*. (WEI)

PRECISIONES TERMINOLÓGICAS Y CONCEPTUALES

El término *indicador* es utilizado en varios sentidos no del todo coincidentes. Para quien está familiarizado con las técnicas de investigación social en la tradición metodológica iniciada por Paul Lazarsfeld, la noción es parte de la conocida trilogía del proceso de operacionalización: variables, dimensiones e indicadores. El concepto clave es el de variable, el cual

denota un aspecto particular de la realidad estudiada, este aspecto puede aislarse de los demás analíticamente, y no permanece constante sino que, precisamente, adopta valores distintos, varía, entre los sujetos de la población a estudiar. Algunas veces las variables se conceptualizan de tal manera que resulta muy sencillo identificar la realidad a la cual corresponden. Es el caso de variables como sexo, edad o estado civil, tan utilizadas en muchas investigaciones.

En otros casos, las variables son conceptualizadas de manera mucho menos clara, de suerte que, para poder utilizarlas en forma práctica, es necesario buscar conceptos que puedan manejarse de manera más precisa y se consideren equivalentes o representativos del concepto más abstracto. Si interesa, por ejemplo, utilizar en un estudio el aspecto (variable) *nivel socioeconómico*, se podrá utilizar en su lugar el concepto más fácil de manejar *ingresos mensuales*. Este último es un *indicador* de la variable anterior. Naturalmente, la traducción de ciertas variables a sus indicadores no necesariamente resulta afortunada.

En forma relacionada con la anterior, pero con algunas diferencias, desde los años sesenta se trató de identificar conceptos particulares precisos que permitieran concretar y valorar otros tan amplios e imprecisos como *desarrollo integral* o *bienestar social*. El concepto de *Producto Interno Bruto*, tan caro a los economistas, fue durante mucho tiempo el único criterio para apreciar el grado de desarrollo de un país. Evidentemente, el concepto ignoraba importantísimas dimensiones de una realidad tan compleja. Por ello se desarrollaron otros *indicadores* de conceptos tan amplios.

La obra colectiva editada por Raymond Bauer con el título *Social Indicators* (1966), fue considerada el *manifiesto del movimiento de los indicadores*, en un momento en el cual la noción se había puesto de moda. En ella puede encontrarse una consideración similar a la que se acaba de hacer, sobre los indicadores como nociones que concretan conceptos más generales, para hacerlos manejables en forma empírica:

El problema clave de un sistema de indicadores sociales... es que nunca podemos medir directamente las variables que nos interesan, sino debemos 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).

Bauer toma de Gross un esquema que ejemplifica el paso de una *gran abstracción* (abundancia) a conceptos más precisos, *indicadores cuantitativos* (en la forma de cantidades de bienes y servicios, producidos en ciertos períodos, y expresados en unidades físicas o monetarias), pasando previamente por *abstracciones intermedias* (riqueza: cantidad, distribución). (Bauer, 1966: 45).

Sin embargo, a diferencia del sentido que adopta en la terminología de Lazarsfeld, en el contexto que ahora nos ocupa la palabra *indicador* no necesariamente se refiere a un as-

pecto particular de una variable más amplia sino que, al contrario, puede tener un carácter sintético, integrando varios más particulares, pero siempre con la pretensión de hacerlo de tal suerte que sea posible un tratamiento preciso, cuantitativo. Así puede entenderse la siguiente definición de indicador: *estadística sintética (summary) sobre el estado actual de un sistema educativo*. (Wyatt 1994: 99).

Según Scheerens los indicadores son estadísticas globales, de baja inferencia, definidas a un elevado nivel de agregación. (1994: 17). Sin embargo, como se verá, también pueden encontrarse indicadores que son todo lo contrario: estadísticas específicas y de alta inferencia derivadas de encuestas de opinión o de investigaciones. Según Wyatt 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; por tanto, el indicador se liga 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).

Wyatt coincide con Oakes en que un indicador debe tener las siguientes características:

- ◆ Medir aspectos que se encuentren en todos los ámbitos del sistema a evaluar (*ubiquitous*), de suerte que puedan hacerse comparaciones.
- ◆ Medir aspectos duraderos del sistema, de suerte que puedan analizarse tendencias y cambios en el tiempo.
- ◆ Ser fácilmente inteligibles para una audiencia amplia.
- ◆ Tener factibilidad, teniendo en cuenta el tiempo, costo y capacidad técnica requeridos para su obtención.
- ◆ Ser generalmente aceptados por sus cualidades técnicas de validez y confiabilidad. (Oakes, 1986: 1-2; Wyatt, 1994:105).

Siguiendo a Jaeger, quien muestra la poca consistencia de las definiciones de indicadores que se encuentran en la literatura especializada (1978), Shavelson y colaboradores consideran conveniente 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).

La noción que adoptó el *Special Study Panel on Indicators* formado en 1989 por el *National Center for Education Statistics* es similar:

Un indicador es una estadística que mide nuestro bienestar colectivo. Un verdadero indicador mide la salud de un sistema, sea económico, de empleo, de servicios médicos o educativos... A diferencia de otras estadísticas, un indicador debe ser relevante para la toma de decisiones, en función de ciertos problemas; debe ofrecer información sobre un rasgo significativo del sistema al que se refiere; y generalmente incluye algún estándar contra el cual pueda juzgarse si hay progreso o retroceso. (*Special Study Panel*, 1991: 12).

Fuera de los Estados Unidos, los esfuerzos por precisar la definición del término indicador han seguido caminos similares, inclinándose por un acercamiento pragmático. Un documento reciente al respecto de la Comisión Europea establece:

Un indicador puede definirse como la medición de un objetivo a perseguir, de un recurso a movilizar, de un efecto alcanzado, de una estimación de calidad, o de una variable de contexto. Un indicador ofrece información cuantificada con el propósito de ayudar a los participantes en acciones públicas a comunicarse, negociar y tomar decisiones... Un indicador cuantifica un elemento que se considera relevante para el monitoreo o la evaluación de un programa. (1999: 17).

Al igual que Shavelson y otros, Oakes distingue dos tipos básicos de indicadores: 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 requerida 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 contenida, 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 ofrecida, 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). La siguiente cita es representativa de estas opiniones:

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).

La noción de *sistema*, junto con las distinciones anteriores entre indicadores elementales y compuestos, agregados o desagregados u otras, permite matizar una recomendación que se encuentra en la literatura, la cual plantea la conveniencia de usar un número reducido de indicadores, ya que un número grande dificultaría la toma de decisiones basada en ellos. En una dirección opuesta, el *Special Study Panel on Education Indicators* del NCES señala:

Pero a medida que nuestro trabajo avanzaba, nos convencimos de que la búsqueda de un número limitado de indicadores educativos clave era equivocada. Como ningún conjunto limitado de indicadores haría justicia de la complejidad de la empresa educativa, un conjunto limitado no sólo reflejaría cierta agenda educativa, sino que definiría una agenda así. Si, por ejemplo, se establece que la enseñanza de las matemáticas y la geografía son suficientemente importantes para merecer que se les dedique un indicador, pero no ocurre lo mismo con la enseñanza de la música o de una lengua extranjera, los educadores tomarán eso como una señal a seguir.

...La fortaleza de los indicadores, en pocas palabras, consiste en que focalizan la atención en temas críticos. Esta cualidad de focalización significa que los indicadores pueden ser palancas para el cambio: pueden transformarse en herramientas para la reforma, porque son instrumentos excelentes para la comunicación. Pero esa misma cualidad es una debilidad poten-

cial: si un conjunto limitado de indicadores focaliza la atención en temas equivocados pueden crearse más problemas educativos de los que se resuelven. (1991: 7).

METODOLOGÍA PARA EL DISEÑO DE UN SISTEMA DE INDICADORES

A partir de la propia experiencia y recogiendo elementos de la literatura revisada (en especial Shavelson *et al.*, 1989: 8-12; y Comisión Europea, 1999: 179-209), a continuación se presenta un intento por sistematizar una metodología para el diseño de sistemas de indicadores.

PRIMERA ETAPA: ELABORACIÓN DEL LISTADO 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); quienes se ubican en instituciones (planteles, agrupados en zonas, ubicadas en municipios y circunscripciones más amplias); en donde hay muebles, equipos y materiales (pupitres, libros de texto, cuadernos, pizarrones, computadoras y muchos más); utilizados en la enseñanza o la administración, con elementos menos tangibles: un currículo, cierta normatividad, algunas tradiciones, etcétera.

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 evaluar la calidad del sistema educativo, precisar la noción de calidad es fundamental para poder 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 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*, 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.
- ◆ *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 así, los objetivos educativos sean alcanzados por el mayor número posible.

Para relacionar los elementos empíricos y las 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*, el cual 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 posibles indicadores, debidamente 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

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

- ◆ *Nombre*: lo más breve y claro 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, señalando la dimensión del concepto de calidad y los elementos sistémicos a los que corresponde: relevancia, eficacia, eficiencia, etcétera; 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, entidad, modalidad de escuela, etcétera).
- ◆ *Elementos para la Interpretación* de los resultados.

TERCERA ETAPA: JUECEO INICIAL

Antes de proceder a organizar costosos procesos de recolección de información para alimentar un sistema de indicadores, cada uno de los elementos que lo forman deberá someterse a un cuidadoso 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, los cuales 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 medir realmente lo que se pretende sea medido.
- ◆ *Confiabilidad*, la otra cualidad técnica esencial, consistente en que el indicador se defina conceptual y operacionalmente de forma tal que, la información obtenida sea consistente a lo largo de las sucesivas aplicaciones del sistema.
- ◆ *Comparabilidad*, si el indicador puede aplicarse 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 ofrecida por el indicador (*freshness*). Uno que no pueda ofrecer información razonablemente reciente no tiene mucho valor.
- ◆ *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 del indicador, en el sentido de que la información necesaria para construirlo 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, el cual tenga peso significativo en relación con alguna dimensión de la calidad.
- ◆ *Utilidad*, como el grado en que el indicador se refiera a un aspecto donde las decisiones tomadas por los responsables puedan impactar de manera significativa.
- ◆ *Claridad*, en el sentido de que la manera como se 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 utilizada 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, etcétera) ofreciendo la información de cada uno que resultó de la segunda etapa.
- ◆ 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, expresar argumentos para defender su posición.
- ◆ En la tercera vuelta 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 de los participantes en el proceso podrán abandonarlo antes de terminar, y el tiempo que implica recoger las opiniones de cada vuelta y procesarlas antes de la siguiente, deberá considerarse un tiempo razonable para cada una (de dos a cuatro semanas) e incluir inicialmente a un número de jueces mayor al que se espera termine el proceso, digamos un mínimo de diez personas de cada una de las tres clases de participantes.

CUARTA ETAPA: PRUEBA PILOTO O PRIMERA APLICACIÓN

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 arrojada por el sistema es interpretada correctamente y utilizada para sustentar acciones de mejora.

ETAPAS POSTERIORES

El sistema podrá operar regularmente, con dos procesos adicionales:

- ◆ Uno de refinamiento permanente de los indicadores que lo requieran, con base en la experiencia que se irá acumulando.
- ◆ Otro de establecimiento y ajuste de parámetros o estándares de referencia, con los cuales se contrastarán los valores de los indicadores, para llegar a un juicio de valor.

No debe olvidarse, en efecto, que la evaluación implica la medición, pero no se reduce a ella. Un indicador captura o mide cierto aspecto o dimensión de la realidad a evaluar, en nuestro caso, del sistema educativo; esto es importante para poder evaluar ese aspecto, pero no es suficiente: para llegar a un juicio sobre la realidad medida hace falta un referente, el cual defina lo que se considera adecuado o inadecuado.

Ese referente con el cual debe contrastarse el resultado de la medición es el *parámetro* o *estándar*. Es normativo, no empírico. Para definirlo, la pregunta a formular es ¿con qué o

con quién debemos compararnos? Para responder la pregunta debemos 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 evalúan. 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 establecer el menor valor aceptable, cuyo cumplimiento permite, por ejemplo, aprobar a un alumno, o acreditar una institución. El uso de los países de menor desarrollo como puntos de comparación podría verse como una aproximación al uso de referentes mínimos.

No es difícil apreciar que cada uno de estos referentes (*óptimo*, *promedio* y *mínimo*) tiene cierto sentido; cada uno arroja luz sobre la realidad evaluada desde cierta perspectiva, y ninguno es suficiente. Parece conveniente usar los tres tipos de parámetro, para lograr una apreciación más 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 plazos. La globalización del mundo contemporáneo lo hace inevitable.

Pero utilizar sólo tales referentes implica una dosis de injusticia, 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, para 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, o una escuela, puede valorar si alcanza o no, si se aproxima o se aleja, a mayor o menor velocidad, a las metas establecidas 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.

LA MADURACIÓN DE LOS SISTEMAS DE INDICADORES

Es importante añadir que, si se quiere que sea sólido, el proceso de desarrollo de indicadores es necesariamente largo.

Para asegurar la solidez, los pasos indicados en las sugerencias metodológicas anteriores deben seguirse cuidadosamente; se debe involucrar a grupos de actores calificados; y los productos del trabajo deben revisarse y corregirse una y otra vez, antes de considerarlos bien establecidos.

En el caso de los indicadores educativos, si se tiene en cuenta el ritmo anual de los procesos de recolección de muchas de las estadísticas básicas obtenidas de las escuelas al inicio y/o fin de cada ciclo escolar, se entenderá mejor el tiempo que implica, no sólo el primer planteamiento de ciertos indicadores, sino su refinamiento hasta alcanzar un grado aceptable de madurez, esto requiere 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 sesenta, 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 aisladamente entre éstos, 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 se debe tener en cuenta que el desarrollo de un sistema de indicadores maduro es una tarea de, al menos, una década.

CONCLUSIÓN

La obra sobre indicadores de la Comisión Europea tiene una reflexión sobre las lecciones de la experiencia internacional, a partir de los trabajos de la especialista recién citada:

*La experiencia internacional en el campo de los indicadores es amplia y ha existido por largo tiempo. En el origen de todos los desarrollos en este dominio encontramos el éxito de unos cuantos 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. Muchas lecciones se sacaron de esta experiencia y se establecieron sólidamente. Pueden resumirse de alguna manera en los títulos de tres publicaciones notables de la misma autora, Judith Innes de Neufville: en 1975 *Social indicators and public policy*; en 1989 *Disappointments and Legacies of Social Indicators*; y en 1994 *Knowledge and Public Policy: the Search for Meaningful Indicators*. (1999: 219).*

De modo similar, Sheldon afirma: *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 haber logrado 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 existentes hasta 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 que se publicó en 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 expresadas por la autora en la segunda edición son importantes para apreciar el cambio de perspectiva producida 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 cual justifica la reimpresión del texto original sin cambios. 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 imperante en la década de 1970, todavía marcada por una concepción ingenuamente positivista. (Innes de Neufville, 2002).

Ésta 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 concepción más compleja de la relación entre conocimiento y acción, a partir de los resultados de sus investigaciones y con apoyo en concepciones más complejas, 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, sino que 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, la cual 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 de construcción mismo, con aportes técnicos, discusiones que aclaren puntos difíciles y consensos más o menos amplios y sólidos laboriosamente alcanzados.

En el campo educativo, un sistema de indicadores no debe ser visto, pues, 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 educativas existentes. Sin olvidar la complejidad de la construcción de consensos a la cual se acaba de hacer referencia, 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 los del 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 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 solo enfoque, bastante estrecho, para monitorear sus sistemas educativos -los conteos de tipo censo- y en un solo 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.

...Aun 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 1991b: 8-9).

Durante 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, los cuales 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 para 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. Pretenden 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, y permitan que transcurran los plazos necesarios para hacerla bien.

Para terminar en un tono más optimista diremos que, para alcanzar una meta ambiciosa, se deben aceptar las limitaciones del punto de partida y mejorar progresivamente; y 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, 1966a: 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

- 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.
- BOTTANI, NORBERT e I. DELFAU (1990). The Search for International Educational Indicators. *The OECD Observer*. Nº 162: 14-18.
- BRYK, ANTHONY y K. L. HERMANSON (1992). Educational Indicator System: Observations on their Structure, Interpretation and Use. *Review of Research in Education*. Vol. 19: 451-484.
- BURSTEIN et al. (1992). Education Indicators. En ALKIN, MARVIN C. Ed. *Encyclopedia of Educational Research*. 6th ed. New York, Macmillan. Vol. 2: 410-418.
- CERI-INES (1995a). *Measuring the Quality of Schools*. Paris. OCDE. ED 411 321, 238 pp.
- CERI-INES (1995b). *Measuring What Students Learn*. Paris. OCDE.
- CERI-INES (1995c). *Public expectations of the final stage of compulsory education*. Paris. OECD.
- CERI-INES (1991a). *An International Handbook of Educational Indicators*. Paris. OCDE.
- CERI-INES (1991b). *General Assembly on International Education Indicators Handbook*. Paris. OCDE.
- CERI-INES (1992). *The OECD International Education Indicators: A Framework for Analysis*. París, OCDE.
- CERI-INES (1973). *Indicators of Performance of Educational Systems*. París, OCDE.
- COMISION EUROPEA (1999). *Selection and use of indicators for monitoring and evaluation. Evaluating socio-economic programmes, Vol. 2*. Luxembourg. Office for Official Publications of the European Communities.
- COMITÉ PRESIDENCIAL DE INVESTIGACIÓN SOBRE TENDENCIAS SOCIALES (1933). *Recent Social Trends in the United States*. New Cork. McGraw Hill.
- DAVID, JANE L. (1988). The Use of Indicators By School Districts. Aid or Threat To Improvement? *Phi Delta Kappan*. March, pp. 499-502.
- DAVID, JANE L. (1987). *Improving Education with Locally Developed Indicators*. New Brunswick. Center for Policy Research in Education-Rutgers University/The Rand Corporation/Wisconsin Center for Education Research.
- DELORS, JACQUES Ed. (1971). *Contributions a une recherche sur les indicateurs sociaux*. París, Futuribles, SEDEIS.
- DEPTO. DE ASUNTOS ECONÓMICOS Y SOCIALES (1975). *Hacia un sistema de estadísticas sociales y demográficas*. Nueva York, ONU.
- (1999). *Educational indicators: an international perspective*. Washington. U. S. Dept. of Education.

EMIN, J. C. (1995). "La mise en place d'un dispositif d'indicateurs pour le pilotage des établissements secondaires français". En CERI-INES 1995a, Cap. 10 pp. 201-231.

FITZ-GIBBON, CAROL TAYLOR (1996). *Monitoring Education: Indicators, Quality and Effectiveness*. London. Cassell.

GILMARTIN, KEVIN J.; ROBERT J. ROSSI; LEONARD S. LUTOMSKI; y DONALD F. B. REED (1979). *Social Indicators: An Annotated Bibliography of Current Literature*. New York-London. Garland Publishing, Inc.

GREANEY, VINCENT y THOMAS KELLAGHAN (1996). *Monitoring the Learning Outcomes of Education Systems*. Washington. World Bank.

GROSS, BERTRAM M. (1966). Preface. A Historical Note on Social Indicators. En BAUER, RAYMOND A. Ed., 1966: ix-xviii.

GROSS, BERTRAM M. (1966). The State of Nation: Social Systems Accounting. En BAUER, RAYMOND A. Ed., 1966: 154-271.

GUILFORD, D. y H. HARTMAN (1981). *Indicators of Equity in Education*. Washington. National Research Council of the National Academy of Sciences.

HOPKINS D. (1994). Process Indicators for School Improvement. En TUIJNMAN-BOTTANI, Cap. 8:145-170.

HOPKINS, D. (1990). International Indicators and Evaluation of Educational Systems. *INES News*, Nº 4: 1-7.

INCE (2000). *Sistema estatal de indicadores de la educación. Síntesis 2000*. Madrid. Ministerio de Educación, Cultura y Deporte.

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*.

INNES DE NEUFVILLE, JUDITH (1989). Dissapointments and Legacies of Social Indicators. *Review of nn*.

INNES DE NEUFVILLE, JUDITH (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.

JAEGER, RICHARD M. (1977). *An abundance of answers in search of questions: On a methodology of assessment through indicators*. Congreso anual de la AERA. ERIC ED 135 872.

JOHNSTONE J. N. (1981). *Indicators of Education Systems*, London, Kogan Page-UNESCO.

KAAGAN S.S. y R.J. LOBY (1989). *State Education Indicators: Measured Strides, Missing Steps* New Brusnswick. Rutgers Univ. Center for Policy Research in Education.

KRATHWOHL, D. R. Ed. (1975). *Educational Indicators : Monitoring the State of Education*. Princeton. Educational Testing Service.

---- (1994). *L 'Etat de l 'ecole*. Paris, Ministere de l'Education Nationale.

LAND, K. C. y S. SPILERMAN Eds. (1975). *Social Indicator Models*. New York. Russell Sage Foundation.

De LANDSHEERE, GILBERT (1994). *Le pilotage des systèmes d'éducation*. Bruselas. De Boeck.

MAGUIRE TH. O. (1993). Some Thoughts on the Alberta Version of Educational Quality Indicators. *The Alberta Journal of Educational Research*. Vol. XXXIX No. 2, pp. 275-277.

MARTÍNEZ RIZO, FELIPE (2002). Revisión de literatura sobre evaluación de sistemas educativos. En VARIOS, *La calidad de la educación en México: Perspectivas, análisis y evaluación*. México. Miguel Ángel Porrúa-SEP-Congreso de la Unión, pp. 305-427.

MARTÍNEZ RIZO, FELIPE (1996). *La calidad de la educación en Aguascalientes. Diseño de un sistema de monitoreo*. Aguascalientes. UAA-IEA.

McEWEN N. (1993). Educational Quality Indicators. *The Alberta Journal of Educational Research*. Vol. XXXIX. No. 2:167-177.

MEYER, ROBERT H. (1992). Can Schools Be Held Accountable for Good Performance? A Critique of Common Educational Performance Indicators. Mecnograma.

MURNANE, RICHARD J. (1987). Improving Education Indicators and Economic Indicators: The Same Problems? *Educational Evaluation and Policy Analysis*. Vol. 9 (2): 101-116.

MURNANE, RICHARD J. y EDWARD W. PAULY (1988). Lessons from Comparing Educational and Economic Indicators. *Phi Delta Kappan*. March, 509-513.

OAKES, JEANNIE (1989). What Educational Indicators? The Case for Assessing the School Context. *Educational Evaluation and Policy Analysis*. Vol. 11 (2), pp. 181-199.

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.

ODDEN, ALLAN (1990). Educational Indicators in the United States: The Need for Analysis. *Educational Researcher*. June-july, pp. 24-29.

OLSON M. (1975). "Measurement and Efficiency in Education". En KRATHWOHL, 1975.

PURYEAR J.M. (1993) "Status and Problems of International Educational Statistics and Research". CIES Newsletter, No. 103:4.

POLLARD, JOYCE S. (1990). *Developping useful educational indicator systems*. Washington. U. S. Dept. of Education.

PORTER, ANDREW C. (1991). Creating a System of School Process Indicators. *Educational Evaluation and Policy Analysis*. Vol. 13 (1): 13-29.

PORTER, ANDREW C. (1988). Indicators: Objective Data Or Political Tool? *Phi Delta Kappan*. March, pp. 503-508.

RICHARDS, CRAIG E. (1988). Indicators and Three Types of Educational Monitoring Systems: Implications for Design. *Phi Delta Kappan*. March, pp. 495-498.

ROSS K. N. y T. N. POSTLETHWAITE (1992). *Indicators of the quality of education: a summa-*

ry of a national study of primary schools in Zimbabwe. Paris, UNESCO-IIEP.

RUMBERGER R.W. (1994). Labour Market Outcomes as Indicators of Educational Performance. En TUIJNMAN-BOTTANI, 1994:265-286.

RUSSETT, BRUCE M., H. R. ALKER, K. W. DEUTSCH y H. D. LASSWELL (1965). *World Handbook of Political and Social Indicators*. New Haven, Yale Univ. Press. Trad. español Euramérica, 1968.

RYLEY, KATHERYN (1999). *Measuring Quality: Education Indicators*. London. Routledge-Falmer.

SALZER, M.; C. T. NIXON; L. J. A. SCHUT; M. S. KARVER ; y L. BICKMAN (1997). Validating Quality Indicators. Quality as Relationship between Structure, Process and Outcome. *Evaluation Review*. 21 (3), 292-309.

SAUVAGEOT C. (Ed.) (1992). *Primary Education in Lesotho. Indicators 1992*. Paris. IIEP.

SAUVAGEOT C. (Ed.) (1993). *L'enseignement fondamental au Mali. Indicateurs 1993*. Paris. IIEP.

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.

SHELDON, ELEANOR BERNERT y ROBERT PARK (1975). Social Indicators. *Science*. 188: 693-699.

SHELDON, ELEANOR BERNERT y H. E. FREEMAN (1971). Notes on Social Indicators: Promises and Potentials. *Policy Sciences*. 1: 97-111.

SHELDON, ELEANOR BERNERT y W. E. MOORE (1968). *Indicators of Social Change*. New York. Russell Sage Foundation.

SHONFIELD A. y S. SHAW, Eds. (1972). *Social Indicators and Social Policy*. London. Heinemann Educational Books.

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. USA Dpt. of Education.

TUIJNMAN A. y N. BOTTANI (Eds.) (1994). *Making Education Count. Developing and Using International Indicators*. Paris. OCDE.

TYMMS, PETER (1999). *Baseline Assessment and Monitoring in Primary Schools. Achievements, Attitudes and Value-added Indicators*. London. David Fulton Publisher.

WICKENS, A. J. (1953). Statistics and the Public Interest. *The Journal of the American Statistical Association*. Vol. 48, pp. 1-14.

WILCOX, LESLIE D.; RALPH M. BROOKS; GEORGE M. BEAL; y GERALD E. KLONGLAN (1972). *Social Indicators and Societal Monitoring. An Annotated Bibliography*. San Francisco. Jossey Bass-Elsevier.

WYATT T. (1994). Education indicators: a review of the literature. En TUIJNMAN y BOTTA-NI, 1994:99-121.