

## EL PAPEL DE LOS VALORES EN LA EVALUACIÓN DE POLÍTICAS PARA SISTEMAS EDUCATIVOS COMPLEJOS

TERESA BRACHO GONZÁLEZ

### Resumen:

El presente artículo propone desarrollar una perspectiva valoral de la evaluación para mejorar los juicios sobre los resultados de la acción pública en contextos de alta complejidad de implementación gubernamental. La discusión se centra en los alcances y dificultades del uso de la Metodología de Marco Lógico y las matrices de indicadores mediante la revisión de 40 evaluaciones de programas educativos federales. Para ello, se acude a dos perspectivas de estudio: los niveles de análisis de las políticas de Fischer y un enfoque de sistemas complejos con base en Morell y en Williams y Hummelbrunner. Esta última concepción permite destacar la relevancia de los agentes involucrados en un programa o una política, la operación bajo contextos cambiantes y la influencia de las disposiciones ajenas a los criterios de la metodología.

### Abstract:

This article proposes the development of a value-based perspective of evaluation in order to improve opinions on the results of public action in highly complex contexts of governmental implementation. The discussion is centered on the scope and difficulties of using the Logical Framework Approach and indicator matrices, based on a review of forty evaluations of federal educational programs. Two perspectives of study are viewed: the levels of analysis of the Fischer policies, and a focus on complex systems based on Morell and Williams and Hummelbrunner. The second conception permits emphasizing the relevance of the agents involved in a program or policy, operation in changing contexts, and the influence of measures foreign to the methodology's criteria.

**Palabras clave:** evaluación de programas, políticas públicas, metodologías, valores y política educativa, México.

**Keywords:** evaluation of programs, public policies, methodologies, values and educational policy, Mexico.

---

Teresa Bracho González es profesora-investigadora de la Facultad Latinoamericana de Ciencias Sociales (FLACSO), sede México. Carretera al Ajusco 377, colonia Héroes de Padierna, 14200, México, DF. CE: teresa.bracho@flacso.edu.mx

La autora agradece las sugerencias de dos jueces anónimos, así como el apoyo de la maestra Mariana Sánchez. Sus comentarios indudablemente contribuyeron al trabajo final.

## Introducción

La evaluación de los resultados de una política educativa, ya sea mediante los rendimientos escolares, la distribución de credenciales en la población o de los bienes y servicios ofrecidos, se entiende como un juicio sobre el mérito o valor de las acciones públicas emprendidas en torno a la organización del sistema educativo, la distribución de oportunidades de aprendizaje para la población, la distribución del conocimiento y certificaciones educativas y sus usos por la sociedad. En el presente artículo, se acota el problema de la evaluación de las acciones gubernamentales al ámbito educativo,<sup>1</sup> no trata en particular ningún programa, política o acción de gobierno. Refiere, sí, al sistema educativo y las políticas públicas que lo regulan y su evaluación.

Toda evaluación de políticas es una práctica de argumentación basada en información adecuada para emitir un juicio sobre las acciones públicas, su pertinencia y sus efectos (Bracho, 2009b). Esto representa importantes retos al evaluador en tanto implica la necesidad de definir, junto con los responsables de la política en cuestión, el valor que orienta la acción gubernamental y su capacidad para obtener los resultados deseados y esperados, así como determinar la clase de evidencia suficiente para emitir una opinión informada sobre la política. De manera muy destacada, la evaluación supone una definición sobre los parámetros con los que se realizarán los juicios, el valor sustantivo de la política para atender a una demanda social, o un valor común, y la forma como habrán de realizarse las apreciaciones sobre los recursos, las actividades y directrices contempladas para brindar una solución a los problemas que se identificaron como atendibles por la acción gubernamental.<sup>2</sup>

En este sentido, la evaluación conlleva atributos que van más allá de la medición y la corroboración de alcance de objetivos en el tiempo con base en ciertos parámetros. La política pública y su evaluación involucran también la adopción de valores y principios sociales que tratan de adecuarse a estructuras metodológicas mediante las cuales se intenta sistematizar la realidad donde se ubica la acción gubernamental y que hace uso de métodos y herramientas para cumplir con ambos cometidos. Estos últimos temas empiezan a destacar en el debate internacional, por lo que juzgo relevante discutirlos para el caso actual de la evaluación de las políticas educativas en el marco de la normatividad reciente sobre evaluación.

Para ello, parto de una revisión del modelo que se ha generalizado en la política pública mexicana: la Metodología de Marco Lógico (MML) y la Matriz de Indicadores de Resultados (MIR) como su expresión empírica. Se discuten las responsabilidades que involucra todo proceso de evaluación de políticas, destacando los siguientes puntos: el papel de los valores públicos en el proceso de evaluación de políticas (compromiso normativo), el papel de la información, su construcción y definición (compromiso empírico), la evaluación de sistemas complejos (compromiso teórico) y, por último, el proceso de comunicación de los resultados de las evaluaciones y su uso (compromiso comunicativo). La argumentación del artículo sigue de cerca una perspectiva de la evaluación desarrollada con base en las ideas de Morell (2010) y de Williams y Hummelbrunner (2009) para destacar los alcances y dificultades metodológicas y de aplicación del marco lógico en el ámbito educativo, incorporando al mismo tiempo la discusión metodológica que distingue diversos niveles analíticos en la evaluación de la política propuesta en Fischer (2006): la técnica-analítica, la contextual, la sistémica y la valoral (como niveles de primer y segundo orden).

Con base en una revisión de las evaluaciones de Diseño, así como de Consistencia y Resultados realizadas a los principales programas de la Secretaría de Educación Pública (SEP) entre 2007 a 2010, se destacan aspectos clave del diseño de los programas, tales como la identificación del problema público, la existencia de diagnósticos actualizados acerca del problema en cuestión, la definición, caracterización y cuantificación de la población objetivo y potencial, así como las fuentes de información de las MIR (en los niveles de fin y propósito) y la participación de las entidades federativas y los municipios en la implementación de los programas. A partir del análisis se busca identificar empíricamente la capacidad que tienen las agencias públicas y los programas de acción gubernamentales para precisar las cuestiones sustantivas que demanda la MML y la construcción de la MIR con el fin de considerar que la metodología se ha utilizado de manera adecuada.

De este modo, en la primera parte del artículo se señalan los principales elementos conceptuales de la herramienta, se hace explícita su utilidad y alcances en el diseño y la evaluación de los programas, se identifican las principales dificultades o limitaciones de la metodología desde un plano normativo y práctico y se destacan los aspectos contextuales más relevantes en su aplicación para el caso nacional y para los programas educativos

considerados. A partir de lo anterior, se señalan algunas implicaciones y obstáculos de la adopción de una herramienta orientada tanto al diseño, como a la evaluación de programas públicos en función de sus resultados, derivados de la formulación e implementación de programas con varios años en operación, así como la presencia de disposiciones externas –i.e. decisiones por parte de agentes políticos no involucrados en su diseño y operación– que inhiben los cambios que demanda la introducción de la MML. A lo anterior se agrega la discusión sobre los problemas más inmediatos de puesta en marcha de la herramienta metodológica para el caso de los programas bajo un contexto descentralizado. Finalmente, se plantea una perspectiva evaluativa que propongo como más adecuada para el caso del sistema educativo mexicano.

### **La Metodología de Marco Lógico: principales conceptos.**

#### **Alcances y dificultades en el diseño y la evaluación**

Numerosos trabajos refieren a la MML, en tanto se trata de una herramienta ampliamente utilizada por organismos internacionales de financiamiento, tales como el Banco Interamericano de Desarrollo (BID), el Banco Mundial, la Comisión Económica para América Latina y el Caribe (CEPAL), entre otros. Se describe aquí la metodología de manera muy breve, con atención particular a la MIR.

Esta metodología supone una identificación de problemas sociales a los que la administración pública trata de responder, precisando cuáles son las lógicas causales que los producen (árbol de problemas), para definir el nivel en que su intervención puede representar una solución adecuada y significativa al problema referido. Supone además una identificación de actores y agencias que pueden estar involucrados en el problema y su resolución (aunque ello no siempre quede expresado en la matriz de indicadores), así como una clara definición de la población en la que se identifica la problemática que el programa buscará atender (véase Ortégón *et al.*, 2005).

Con base en el árbol de problemas que sintetiza las causas (identificadas en el diagnóstico) se construye la Matriz de Indicadores de Resultados en donde se traducen los diversos niveles de objetivos (fin, propósito, componentes y actividades), y la definición de los indicadores que reflejarán el alcance de metas cuando el programa esté en operación, lo que representa una base para su monitoreo y seguimiento.

El Modelo de Marco Lógico, con base en el cual se sintetizan los diseños de los programas y políticas del gobierno federal, y cuyos indicadores dan cuenta del desempeño logrado por éstos [consúltense la página del Consejo Nacional de Evaluación de la Política de Desarrollo Social (Coneval, 2010)] utiliza una matriz de cuatro columnas por cuatro renglones, que se expresa en el cuadro 1.

CUADRO 1  
*La matriz de indicadores de Marco Lógico*

	Objetivos	Indicadores	Medios de verificación	Supuestos
Fin				
Propósito				
Componentes				
Actividades				

Construcción propia.

La identificación de la lógica causal entre las actividades propuestas por el programa público hasta la resolución del problema, se expresan en lo que se denomina Resumen Narrativo de la matriz de indicadores, y que corresponde a la columna de “objetivos”. El cuadro 2 resume los elementos que lo integran y la verificación de la lógica causal que corresponde a la coherencia desde las actividades hasta el alcance del fin. De esta manera las actividades (acciones que habrán de realizarse por la acción pública propuesta) han de conducir a la realización de los componentes (como los servicios ofrecidos por el programa o la acción pública); éstos han de ser capaces de alcanzar el propósito (como el objetivo inmediato de la política) que a su vez conduce a la consecución del fin (como el objetivo general al que la política aporta para la resolución del problema identificado).

CUADRO 2  
*La matriz de indicadores de marco lógico. Objetivos*

	<b>OBJETIVOS</b> (Resumen narrativo)
<b>FIN</b> ↑	Objetivo estratégico en que se inserta Cómo contribuye el programa a solucionar un problema público mayor
<b>PROPÓSITO</b> ↑	Objetivo del programa, <b>lo que se espera lograr</b> Impacto o resultado directo Qué se espera a partir de la realización de los componentes <b>Es la hipótesis central del proyecto</b>
<b>COMPONENTES</b> ↑	<b>Bienes, servicio o infraestructura</b> que se producen con la intención de lograr el objetivo central <b>Bien o servicio que se ofrece a la población</b>
<b>ACTIVIDADES</b> ↑	<b>Cómo se producen los bienes y servicios.</b> Lo que el ejecutor realiza para producir cada componente. Implican el uso de recursos. Se especifican para cada componente.

↑  
Lógica vertical

Construcción propia.

De otro lado, la lógica horizontal de la matriz de indicadores constituye la forma para verificar la operatividad del fin, el propósito, los componentes y las actividades (cuadro 3). Las columnas representan el objetivo del programa, los indicadores que permiten realizar el monitoreo de los avances y las fuentes o medios de verificación en que está contenida la información para la estimación de los indicadores. La última columna representa los supuestos externos para el alcance de los objetivos, expresando de esta manera los riesgos del programa.

El modelo lógico permite responder a tres preguntas fundamentales en el diseño y la evaluación de los programas: por qué, para qué y cómo, lo cual representa de manera sustancial los fines y propósitos esperados de la acción pública. Para contestar estas preguntas se requiere articular recursos y actividades a resultados específicos dirigidos a una población objetivo. De acuerdo con estos parámetros, podemos identificar los principales objetivos, así como el contexto en que se desarrolla el programa; sin embargo, no es posible identificar las relaciones entre los involucrados en el nivel de la entrega del bien o servicio y las formas organizacionales que se dan entre ellos, aspectos propios de la teoría de proceso del programa (Weiss, 1997 y Rossi, *et al.*, 2004)<sup>3</sup> y no consideradas en la MML.

CUADRO 3

*La matriz de indicadores de marco lógico. Indicadores y medios de verificación*

	OBJETIVO	INDICADORES	MEDIOS DE VERIFICACIÓN	SUPUESTO
Fin, Propósito, Componentes y actividades		Medición de objetivos. Permiten monitorerar el avance de los objetivos	Dónde se obtiene la información necesaria para medición de indicadores. Incluye temporalidad, frecuencia de la medición	
<div style="border: 1px solid black; border-radius: 15px; padding: 5px; background-color: #cccccc;">                     Seguimiento y control del programa                      Sustrato empírico de la evaluación                 </div>				



Construcción propia.

Como parte de un sistema de evaluación, la matriz de indicadores, provee una de sus máximas exigencias: un parámetro concreto y medible, justificado con base en una lógica causal, que indica en qué medida se están logrando los objetivos planteados como alcanzables en la política. Entre sus virtudes pueden señalarse la simplicidad que permite entender y comparar políticas generadas en distintos espacios de la acción pública.<sup>4</sup> Permite, además, reconocer la lógica causal que supone la intervención pública concreta, sin requerir de un mayor estudio sobre ese programa. Su expresión en un esquema uniforme y común agrega una formalización de la lógica de las intervenciones públicas, identificando posibles traslapes y metas comunes entre distintas acciones de gobierno.

Es claro que las herramientas técnico-analíticas, como es el caso de la metodología particular que analizo aquí, resuelven importantes aspectos de la función empírica de la evaluación. De acuerdo con Gasper (2000), entre las ventajas de la MML caben destacarse las siguientes: provee una visión global sobre los objetivos de una política o programa; facilita el pensamiento lógico (pero no lo sustituye); permite construir y dar segui-

miento a un conjunto de indicadores de logro, medibles y con plazos de realización de los objetivos; identifica fuentes de información específicas, válidas y comprobables; y permite delimitar algunas hipótesis sobre otros factores incidentes en el programa (pero que quedan en lo previsible como riesgos, al estar capturados en la columna de supuestos).

De otro lado, el marco lógico se centra en efectos y resultados esperados mediante rutas previstas; el logro de los objetivos planeados puede resultar problemático cuando intervienen diversas organizaciones o dependencias, su aplicación a partir de un formato único (matriz de cuatro columnas y cuatro renglones) puede no coincidir con la lógica del diseño en programas preexistentes o el bloqueo de aspectos relevantes (Gasper, 2000). Lo anterior no sólo tiene repercusiones negativas en el ámbito normativo y empírico, sino que también resulta contraproducente cuando se trata de evaluación de sistemas complejos, como se detalla más adelante.

En este sentido, los principales problemas que pueden presentarse en la aplicación de la herramienta (Bracho, 2009a) tienen que ver con:

- su adaptación a programas con varios años de operación y con rutinas técnico-administrativas consolidadas;
- el uso de teorías causales inapropiadas, incorrectas o incompletas, simplificaciones excesivas de la lógica causal o del contexto del programa;
- el reducido margen para identificar los cambios entre las relaciones de los participantes del programa a través de las diferentes etapas de la política, y en algunos casos por la omisión de alternativas de diseños lógicos disponibles;<sup>5</sup>
- la ausencia o falta de información suficiente para la definición de indicadores pertinentes; y,
- las situaciones que no pueden ser previstas ni capturadas en la matriz, referidas a los cambios en el tiempo, la complejidad de sistemas interactivos, los distintos órdenes de gobierno que intervienen en una acción pública, etcétera.

### **El contexto de uso de la MML y la MIR en México**

El uso de la MML y la construcción de las MIR en la Administración Pública Federal (APF), inició de manera generalizada hacia finales del primer trimestre de 2007, como respuesta a las disposiciones de los “Lineamientos generales para la evaluación de programas federales” –emitidos de manera

conjunta por las secretarías de Hacienda y Crédito Público (SHCP) y de la Función Pública (SFP) y el Coneval– y el Presupuesto de Egresos de la Federación del mismo año. Asimismo, con la publicación del Programa Anual de Evaluación (PAE), se introdujeron los criterios técnicos para regular la utilidad y calidad de la acción gubernamental y se dieron a conocer el tipo de evaluaciones a realizar así como los programas federales que habrían de llevarlas a cabo (Castro *et al.*, 2009; González, 2008 y Sánchez, 2011).

Según Sánchez (2011) y con base en entrevistas con los principales actores participantes en el proceso por parte de Coneval, entre las razones que motivaron el uso de la metodología en la APF se encuentran: una herramienta de fácil explicación que aportaba un lenguaje común orientado a resultados; la construcción de indicadores que permitiría sentar las bases para erigir un sistema de monitoreo y evaluación en las diversas dependencias, la experiencia chilena como un referente internacional en donde se había ya generalizado el uso de la herramienta en la acción gubernamental, así como la prueba piloto realizada en la Secretaría de Desarrollo Social apenas tres años antes. Estos eventos avalaron su utilización, en una primera instancia, para los programas federales que operaban con reglas de operación (ROP).

A partir de este contexto, la MML como herramienta de planificación, representa de forma sistemática y ordenada los objetivos de un proyecto o programa a partir de relaciones de causalidad entre los diferentes niveles de objetivos, los insumos previstos, las actividades planeadas y los resultados esperados, y define las condiciones fuera del control de la gerencia del programa que pueden influir en su éxito (Agencia de Noruega para la Cooperación para el Desarrollo, NORAD, 1993 y Comisión Europea, 2001 en Bracho 2009a). Con base en las ideas anteriores, la metodología ha servido para mejorar el diseño de los programas, reorientando acciones y beneficios que responden a la identificación de causas de los problemas públicos.

De otro lado, la MML conlleva dos importantes funciones para la evaluación: 1) es un parámetro, criterio o estándar para la examinar el diseño de los programas; y 2) es la fuente de información sobre su desempeño. Estas tareas se ven reflejadas en el Sistema de Evaluación del Desempeño (SED). La matriz de marco lógico, como referencia de evaluación implica una congruencia entre los principales elementos que constituyen la formulación de un programa (Sánchez, 2011):

- 1) la adopción de una perspectiva teórico-conceptual en que se basa la intervención, a fin de trazar la cadena de objetivos con base en una estructura causal;
- 2) la población objetivo que refiere a los atributos sociales, económicos y físicos de las personas a quienes deben orientarse los bienes y servicios de la intervención;
- 3) la selección (distinción) de los procesos operativos que buscan la optimización de los recursos del programa y que además son congruentes con los objetivos que se persiguen; y,
- 4) la identificación de los involucrados en el problema público.

Es importante insistir en la calidad del diseño del programa, sintetizado en la matriz de indicadores, pues de ella dependerá la calidad y coherencia de la información de desempeño expresada en los indicadores. A partir de ellos es posible evaluar el avance en el cumplimiento de los objetivos y dar seguimiento a su operación. De acuerdo con los “Lineamientos” emitidos en 2007, el fin y los objetivos de los programas federales contribuyen al cumplimiento de objetivos estratégicos de dependencias y entidades, los cuales son medidos a través de indicadores de resultados de las MIR; al mismo tiempo, estos objetivos tienen como referente un proceso presupuestario (artículos 5,6 y 7).

Con base en lo anterior, la MIR forma parte del SED para coadyuvar en la eficiencia, economía, eficacia y calidad de la APF, así como el impacto social del ejercicio del gasto público. Es por ello que las dependencias y entidades deberán reportar el avance y resultado de los indicadores incluidos en dicha matriz (artículo 14). De acuerdo con lo anterior, la MML no es sólo la vía para re-analizar los principales elementos de diseño de los programas; trata también de determinar en qué medida lo planificado se está llevando a cabo, y cuáles son los resultados de la acción, cuáles fueron los recursos ejercidos y qué cambios produjo en los beneficiarios.

En la evaluación propuesta por la normatividad vigente, los beneficios de la MIR se extienden hacia el resto de las evaluaciones que conforman el SED de la siguiente manera (Sánchez, 2011): en la evaluación de consistencia y resultados, la matriz de indicadores es la fuente de información primaria de los elementos de diseño del programa para examinar su congruencia según una estructura causal; en la evaluación específica de desempeño 2008-2009, la descripción del diseño de los programas y una primera ex-

ploración sobre el avance de los resultados, se realiza con base en la matriz de indicadores; en la evaluación de procesos, la matriz de marco lógico se convierte en una guía elemental en tanto que identifica las actividades más relevantes del programa. En el caso de la evaluación de impacto, la MIR es sólo una vía para comprender el diseño del programa de manera sintética; sin embargo, la explicación de efectos requiere de información y estudios adicionales y, por cuestiones de temporalidad y complejidad, la información que proveen los indicadores a nivel de propósito o fin es insuficiente. La contribución de la matriz de indicadores a niveles sectorial o nacional se ubica mediante las evaluaciones estratégicas, pues ayuda a discernir y comparar si se están llevando a cabo programas con objetivos y metas comunes, y facilitar su comparación, dada la estructura común de las MIR. Finalmente, en el caso de las evaluaciones específicas, la utilidad de la matriz de indicadores es ambigua, en tanto que se trata de propuestas *ad hoc* reguladas sólo de manera general y sin especificar el uso de alguna metodología específica (artículo 17).

Respecto de la evidencia empírica, hasta el momento son escasos los estudios sobre la aplicación y uso de la MML en la APF; recientemente, la investigación de Sánchez (2011) muestra la relevancia de considerar las condiciones en que se llevó a cabo la aplicación de la metodología y la construcción de la MIR, y que bien pueden identificarse en la mayoría de los programas públicos:

- 1) El uso *ex post*. La construcción de la matriz de indicadores se realizó a partir de un diseño previo, que no consideraba la MML, y en condiciones de una lógica de intervención y operativa consolidada. Los programas pretendían cumplir un sinnúmero de objetivos, buscaban la eficiencia operativa y su diseño estaba sujeto a una normatividad (las ROP) y a una estructura que no era la de metodología de marco lógico, y no siempre era consistente con ella.
- 2) La utilización incompleta de la MML. En general, no se realizó un ejercicio de definición del problema, objetivos ni estrategias mediante los elementos de la metodología; tampoco se realizaron diagnósticos o estudios para apoyar la lógica causal expresada en las MIR. La construcción de árbol de problemas se realizó de manera simplista, generalmente a partir de la lógica previa del programa, más que como una estrategia diagnóstica.

- 3) La presencia de órdenes externas, es decir, decisiones emitidas por actores ajenos al diseño y manejo del programa (como son funcionarios de alto nivel o autoridades políticas), que subestimaron los criterios de la MML, específicamente su estructura causal, lo cual debilitó los alcances de la herramienta o, incluso, llegó a contradecirla.
- 4) La aplicación de la MIR bajo formato único (cuatro columnas y cuatro renglones) inhibió las adecuaciones a programas que establecían la participación y coordinación de los tres niveles de gobierno, donde el gerente y el ejecutor son actores o entidades gubernamentales diferentes; y, por otra parte, desestimó la existencia de diseños complejos cuya lógica de intervención requería una secuencia lógica y temporal de componentes.

A partir de estos aspectos Sánchez concluye que, en lo que respecta al diseño, la MML y la MIR contribuyeron a identificar y reformular las dimensiones del problema social que trata cada programa; sin embargo, la incongruencia en la formulación de la intervención se debió a la presencia de órdenes externas que subestimaron la estructura causal implícita en la MML, limitando así las mejoras en el diseño del programa.<sup>6</sup> Asimismo, la aplicación de la metodología tuvo dificultad para captar la oferta de bienes y servicios que provee el programa y su interacción o recepción por parte de la población objetivo.

#### **Alcances y dificultades tras la adopción de la MML y la construcción de MIR en el ámbito educativo**

Se consideran 40 evaluaciones de programas educativos, las cuales, en su mayoría, corresponden a la modalidad “s”, “Programas sujetos a reglas de operación”. Adicionalmente, se incluyen ocho programas “u”, “b” y “e”, “Otros programas de subsidios”, “Provisión de bienes públicos” y “Prestación de servicios públicos”, respectivamente. Estas evaluaciones de Diseño y Consistencia y Resultados se realizaron entre 2007 y 2010.<sup>7</sup>

Cabe señalar que únicamente 10 de los 40 programas se introdujeron después de 2007, año en que comenzó a utilizarse la MML y la construcción de MIR.<sup>8</sup> Este hecho nos permite afirmar que la herramienta se aplicó de manera *ex post*, y con base en ello destacamos los siguientes aspectos:

- La mayoría de los programas identifica el problema que atiende; los que no lo logran (10) son de reciente operación (2008 y 2009) y corresponden a alguna de las modalidades “u”, “b” y “e”.<sup>9</sup>

- Sólo la mitad de los programas contaba con un diagnóstico actualizado acerca del problema que atienden. De la otra mitad, sólo ocho iniciaron operaciones después de 2007 y 13 de ellos están sujetos a ROP.
- Del total, tres programas son operados por instituciones públicas de educación superior o media superior; 13 operan sólo en las entidades federativas, mientras que 18 involucran tanto a las entidades federativas como al Distrito Federal y cuatro tienen la particularidad de ejecutarse también en los municipios.
- Respecto de la definición, caracterización y cuantificación de las poblaciones objetivo y potencial, sólo doce programas abarcan los tres aspectos. De ellos, los bienes y servicios se proveen en las diferentes entidades federativas y el Distrito Federal y, en solo un caso no está dentro de la modalidad “s” e intervienen diversas instituciones públicas de educación superior.<sup>10</sup> Por el contrario, once programas no definen, caracterizan ni cuantifican sus poblaciones objetivo y potencial; en este subconjunto se ubican dos de los orientados a instituciones de educación superior; el resto (17) cumple parcialmente con dichos aspectos, en su mayoría, definen la población objetivo, pero no la caracterizan ni la cuantifican; 16 de ellos están sujetos a ROP y la mayoría son coordinados por la Subsecretaría de Educación Básica.
- De los 40 programas evaluados, 24 utilizan una fuente externa para la verificación en el nivel de fin; 14 usan una fuente interna para dicho nivel. En seis se utilizan los resultados de la prueba Evaluación Nacional del Logro Académico en Centros Escolares (Enlace) como medio de verificación para fin y propósito y cinco programas abarcan ambos aspectos (un medio de verificación externo y la prueba Enlace).<sup>11</sup>
- Respecto de la frecuencia de medición de los indicadores encontramos que, con la información disponible,<sup>12</sup> 4 programas de educación básica y de reciente creación (2007), cuentan con ROP y miden sexenalmente su fin; a diferencia de 15 que plantean la medición de su fin de forma anual, la mayoría son de modalidad “s” y tienen presencia en las entidades federativas y el Distrito Federal. El resto (5), proponen medir su fin cada dos, tres o cinco años.

Las implicaciones tras el uso de la MML y la construcción de MIR en el diseño y la evaluación de los programas pueden considerarse desde el punto de

vista normativo o técnico de la herramienta y desde una perspectiva práctica relacionada, invariablemente, con el enfoque a resultados. De acuerdo con la revisión de los principales programas educativos se identificó que el uso de la herramienta se ha extendido a toda la acción gubernamental; como se mencionó antes, se aplicaba sólo a los programas modalidad “s”, mientras que a cuatro años de la publicación de los “Lineamientos” de 2007, se ha realizado un esfuerzo para abarcar las modalidades “u”, “b” y “e”. En este caso pueden identificarse importantes deficiencias a nivel de definición del problema y la falta de diagnósticos lo que indica un desapego metodológico en el diseño de la política educativa, hecho que impactará negativamente en la obtención y medición de resultados. Por otra parte, hay que señalar que el uso de la MML y la construcción de MIR se llevan a cabo en programas cuya gestión y operación se ha ido consolidando a través del tiempo (por ejemplo, algunos de los aquí considerados operan desde 1959).

Esta revisión del estado que guarda el diseño de programas educativos desde la perspectiva de la evaluación permite identificar la importante presencia de las entidades federativas, los municipios así como las instituciones de educación superior y media superior para el cumplimiento de objetivos nacionales, pero que difícilmente pueden ser capturados en la lógica del diseño de la MML. En este sentido, cabe recalcar que entre los aspectos problemáticos a considerar se encuentran no sólo la formulación adecuada de los procesos para la implementación en los que interviene tanto la descentralización como las relaciones intergubernamentales,<sup>13</sup> sino también el diseño de estrategias metodológicas para avanzar en la caracterización y la cuantificación de las poblaciones objetivo y potencial, sobre todo en los casos que éstas no se pueden conjuntar (en particular, en algunos programas u, b, e).

El análisis de Bracho *et al.* (2010) muestra que la MML y la construcción de MIR bajo formato único enfrenta serias dificultades cuando se aplica a programas que integran diferentes unidades responsables, que además son de distintos tamaños, que atienden poblaciones potencial y objetivo que no pueden agruparse, o bien, en los casos donde no hay una entidad gubernamental responsable de la operación, y las actividades no refieren a los mismos momentos. Estos elementos clave del diseño de los programas tienen repercusiones relevantes en el nivel del seguimiento y la medición de resultados. En estas condiciones, la MIR no es capaz de reflejar que se conforma por sub-programas que operan de manera autónoma, con objetivos

y poblaciones potencial y objetivo propios; es decir, no se identifican de manera clara los servicios que el programa provee para el fin establecido. La presencia de varias unidades o entidades responsables tiene como consecuencia la falta de expresión de los servicios que ofrece cada institución.<sup>14</sup>

A pesar de estas complicaciones, Bracho *et al.* (2010) proponen como alternativa la matriz anidada o en cascada que permite agregar conceptualmente la interrelación de los distintos sub-programas (sean apoyos o instituciones) orientados a un fin común, aun cuando no sean parte de una agrupación real en la administración, ni tengan una lógica de operación común. Esta propuesta representa una vía para expresar adecuadamente las diversas lógicas de operación que pueden integrar un solo programa, podría mejorar la definición y cuantificación de las poblaciones objetivo y potencial e, incluso, los procesos de implementación para los municipios, entidades federativas así como las distintas unidades educativas responsables que operan bajo un programa.

Por otra parte, el seguimiento de resultados a través de los indicadores no es un asunto menor, donde, evidentemente la calidad del Resumen narrativo condiciona la generación de indicadores adecuados y pertinentes. A ello se suma el vínculo de la MIR con la toma de decisión en el ámbito presupuestal que se refleja en la tendencia de medir sexenalmente el logro del fin, como sucede en 4 de los programas aquí considerados, aspecto que debería estar ligado a la producción de resultados de acuerdo a una lógica causal, no al término de un periodo presidencial.

Con base en lo anterior, de acuerdo con la lógica de implantación de los modelos de evaluación de las políticas mexicanas, el problema se agudiza en tanto que la construcción de las matrices de indicadores supusieron precisiones impuestas desde la autoridad, por ejemplo, tomar como fin el logro educativo además de haber sido construidas sobre programas e instituciones que estaban ya funcionando con otras lógicas de diseño y operación, con base en diagnósticos débiles y poco articulados entre ellos, entre otros problemas. Asimismo, se trató en el caso mexicano de una aplicación incompleta y *ex post* de la MML pues el esfuerzo se centró en obtener una MIR “razonable” en periodos reducidos (sin importar la complejidad de cada programa o política, o la lógica de operación de instituciones ya consolidadas bajo otros esquemas de planeación) y bajo la observancia de instituciones como la SHCP que vincula la información de sus indicadores a procesos programáticos y presupuestales.<sup>15</sup>

Por último, las implicaciones normativas y empíricas derivadas de las limitaciones del marco lógico son relevantes porque de ellas depende la generación de información para el avance del conocimiento en la materia, la toma de decisiones, incluidos los sistemas de monitoreo y evaluación del desempeño, y el mejoramiento de las políticas. No sobra señalar que en el caso nacional, el sistema de evaluación se plantea como una herramienta en la presupuestación del gobierno federal y la posibilidad de tomar decisiones sobre la terminación o continuación de los programas sujetos a evaluación. Este hecho inclina la balanza en contra de posibles esfuerzos teóricos, empíricos y comunicativos, y aleja el diseño y evaluación de criterios orientados a valores y principios nacionales en favor de la educación.

En síntesis, la medición y seguimiento del logro educativo a través de la definición de objetivos congruentes según un conjunto de valores y resultados de amplio alcance queda fuera de la mira de ejecutores, diseñadores y directores de los programas. Los esfuerzos parecen orientarse al cumplimiento de los esquemas programáticos y presupuestales y, en segundo plano, a la asimilación y apropiación de una herramienta metodológica con bases causales.

Este aspecto nos permite discutir los avances de la evaluación educativa en México de acuerdo a lo que identifiqué como los compromisos de la evaluación, los cuales se exponen a continuación. A partir de ellos pretendo destacar un enfoque integral y de amplio espectro que involucra esfuerzos teóricos que, en los hechos, no contradice la adopción de la MML; compromisos empíricos y comunicativos, que son la esencia de la toma de decisiones basada en la evidencia, y el compromiso normativo, aún incipiente dentro del sistema de evaluación nacional y educativo, cuya consideración puede favorecer y mejorar tanto los esquemas de evaluación existentes como la acción pública en general en un contexto de complejidad e incertidumbre que se explica posteriormente.

### **Explicitando los compromisos de la evaluación**

La evaluación de programas y políticas involucra la deliberación de cuestiones tanto normativas como empíricas. La evaluación, al igual que todo proceso investigativo, busca generar conocimiento sobre las políticas y sus resultados mediante el uso de métodos científicos de observación, registro y análisis; sin embargo, se distingue porque busca explícitamente que sus

resultados sean utilizados como insumo o evidencia de un proceso de toma de decisiones que bien puede culminar en la mejora de algunas de las etapas de una política, o bien en su terminación.

De manera destacada, la evaluación trabaja a través de juicios que toman como base valores. Por ello, la elección, definición y consenso de dichos valores representan en todo momento los límites donde se lleva a cabo la deliberación sobre la acción pública. Así, este sustrato valoral del proceso de evaluación de las políticas es, en definitiva, uno de sus compromisos iniciales que marcan todo el proceso investigativo. El compromiso normativo de la evaluación refiere fundamentalmente a la consideración de la legitimidad de los valores sobre los cuales se construyen las decisiones públicas y que constituyen los principales parámetros de juicio en el proceso evaluativo.

El compromiso teórico representa un acotamiento de la perspectiva con base en la cual se decide sobre los acercamientos, métodos e información relevante en cada contexto investigativo. Se trata de un compromiso fundamental de la evaluación que representa su capacidad de comprender teóricamente tanto la problemática social de que se ocupa la acción pública en cuestión, como la acción pública en sí misma. Es decir, representa tanto una teoría sobre el problema social, o una forma particular de entender y concebir teóricamente el problema que atiende la política, como una teoría sobre las políticas públicas y su operación, o sea, una concepción de la intervención pública en los problemas sociales.

El compromiso empírico de la evaluación, al igual que la etapa previa de definición de los problemas públicos y la construcción de diseños de programas y políticas, se enfrenta a una amplia gama de perspectivas conceptuales, la inclusión (o no) de múltiples disciplinas, así como muy variadas alternativas metodológicas, la recuperación de información y la construcción de evidencia empírica, etcétera.

Una vez elaborado un marco de argumentación, la información adquiere un valor diferenciado en función de los usuarios potenciales de los resultados de la evaluación. Habrá de ser explícito el o los usos de la información que se definió como pertinente, sus interpretaciones válidas y los usuarios de los resultados de la evaluación. Con ello, se refiere al último compromiso de la evaluación, el comunicativo. Dicho brevemente, no es lo mismo el uso de la información producida para la toma de decisiones

en el ámbito de las políticas, que su uso para la investigación científica, para la divulgación o para su utilización en el aula (Tiana, 1997; Martínez Rizo y Santos, 2009).

En suma, destaco aquí que un primer acotamiento en todo proceso de evaluación de acciones públicas radica en identificar la definición y el diseño de los problemas públicos que elige un gobierno particular para atender, así como las acciones que juzga pertinentes para su resolución o los marcos de intervención que elige para la resolución de los problemas identificados. Del mismo modo, la definición y selección de teorías sustantivas sobre el problema en cuestión, son un segundo acotamiento al proceso evaluativo al que se agregan las concepciones que se sostienen sobre la acción pública. Finalmente, la definición de la evidencia sobre la cual se emitirán dichos juicios conforma el marco en que será posible emitir un juicio sustantivo sobre la acción pública a que refiere la evaluación.

Partiendo de estos compromisos de la evaluación, es necesario plantearse cómo abordar los problemas de valoración del objeto de la política y de la clase de teorías que son capaces de abordar una explicación plausible del problema en cuestión. Sólo a partir de ello puede analizarse el tema de la pertinencia de la acción pública propuesta para la resolución del problema específico en el sistema social en que se identifica y, por tanto, para poder plantear la evaluación referida al modelo causal de intervención pública para la resolución de problemas sociales.

### **Otras perspectivas de análisis en la evaluación**

Sin pretender cubrir en este espacio toda la literatura sobre evaluación de las políticas públicas, destaco aquí dos perspectivas que, a mi juicio, han de ser relevantes para discutir los modelos que predominan en este momento en el ámbito nacional. Las juzgo importantes por referir a perspectivas un tanto más amplias de la acción pública, que incluyen la relevancia de los valores y la perspectiva de sistemas complejos en los procesos evaluativos.<sup>16</sup> De ahí que me limite a sólo dos propuestas para la comprensión de los *modelos* de evaluación.

La política pública es una acción compleja en sí misma, por lo que limitar su evaluación a sus aspectos técnicos simples no hace más que caricaturizarla. Siguiendo a Fischer (2006), la política pública puede analizarse según cuatro dimensiones: técnica- analítica, contextual, sistémica y valoral. El cuadro 4 resume estas dimensiones.

CUADRO 4

*Lógica de la evaluación de políticas. Niveles, discursos y preguntas*

Nivel	Discurso	Tipo	Preguntas
Primer orden	Técnico analítico	Verificación (outcomes)	El programa, ¿alcanza sus objetivos explícitos?
	Contextual	Validación situacional (objetivos)	El objetivo(s) del programa, ¿es relevante a la situación identificada en el problema que atiende?
Segundo orden	Sistémico	Justificación social (metas)	Los fines del programa, ¿tiene un valor instrumental o contribuyen a la sociedad en su conjunto?
	Valoral	Decisiones sociales (valores)	Los ideales que organizan el orden social, ¿proveen de una base legítima para la resolución de juicios en conflicto?

Fuente: Construcción propia a partir de Fischer (2006).

Las evaluaciones de primer orden parecieran aproximarse a los objetivos del modelo de marco lógico al que he referido en secciones anteriores. En este punto, el interés de la evaluación es saber si la política o el programa público está produciendo los bienes y servicios que definió como relevantes, si alcanza sus metas explícitas y si se enmarca en un diagnóstico que indica la pertinencia o no de dicha política. Incluye pues, el nivel técnico-analítico, el cual consiste en la identificación o verificación del grado en que los objetivos de una política están siendo alcanzados. Asimismo, el nivel contextual permite determinar la relevancia de la acción pública para atender un problema específico y busca explicar si el programa está contribuyendo a la situación considerada como problemática o indeseable, y para la cual se decidió la intervención pública.

En el segundo orden, se sitúan la dimensión sistémica y la valoral. La primera se ocupa de las metas del programa que permiten justificar socialmente la política o programa en cuestión por su valor instrumental o el aporte social. En cambio, la dimensión valoral permite indagar en qué medida los ideales que son propios al orden social proveen una base legítima para resolución de juicios en conflicto.

De esta manera, en las evaluaciones del primer orden se indaga sobre la parte micro de la política (los individuos, las acciones y servicios involucrados en la acción pública), mientras que el segundo se refiere a los aspectos macro (las instituciones en que se inserta el programa o acción

y un sistema social que otorga sentido a las acciones públicas). Por otra parte, cada uno de los niveles permite plantear distintos tipos de preguntas, aunque complementarias entre sí. Cada tipo de evaluación ha de cumplir con los compromisos normativo, teórico y empírico referidos antes. En todo caso hay que diferenciar las fases de la política, sea de diseño, de implementación u operación del programa; en esta última deben considerarse los mecanismos emergentes entre los involucrados y donde el aprendizaje y la evolución de dichos agentes generan sistemas complejos que permitirían capturar el entramado de las acciones públicas en su contexto; por ejemplo, en el caso de sistemas de operación descentralizados; o de cambios de las acciones planteadas originalmente, en su traducción y adaptación en el tiempo y en los distintos niveles de la administración pública.

El paso a través de los cuatro niveles de análisis me ayuda para mostrar que tanto la función normativa como la empírica de la política son partes complementarias y no excluyentes. Su uso integral permitiría transitar de la verificación de resultados y la validación situacional (nivel máximo al que puede aspirar la MML), a una justificación social de la acción gubernamental para la construcción o fortalecimiento de determinado orden en la sociedad.<sup>17</sup>

Este planteamiento es consistente con una perspectiva sistémica de la evaluación de políticas sociales que, a su vez, se complementa con los intentos recientes por identificar en los modelos de evaluación la existencia de aspectos poco previsibles en el diseño de la política, que pueden ser más o menos identificados mediante una investigación que tome como base el reconocimiento de que las políticas públicas operan en el marco de sistemas sociales complejos. Sintetizo para ello algunos preceptos fundamentales (con base en Morell, 2010):

- 1) Los sistemas complejos muestran relaciones no lineales entre sus elementos. Por ello, una pequeña perturbación puede causar un gran efecto, un efecto proporcional o, incluso, ningún efecto en lo absoluto.
- 2) Los sistemas complejos tienen la capacidad de atenuar o enfrentarse a lo imprevisible, pueden retroalimentarse y además tienen la capacidad de aprender y evolucionar con el tiempo.
- 3) Los sistemas complejos se pueden entrelazar; dicho de otro modo, los componentes de estos sistemas pueden integrarse de otros también complejos.

- 4) Los límites de todo sistema, pero en particular de los complejos, son difíciles de determinar.
- 5) Los fenómenos emergentes puedan estar presentes. La suma de comportamientos o acciones no reflejará necesariamente un patrón de acción o tendencia.

A la luz de una perspectiva sistémica, las evaluaciones de primer orden, es decir, dos primeros niveles de análisis de la política, el técnico-analítico y el contextual, permiten abarcar aspectos generales de la evaluación. La tendencia a producir información sobre los resultados alcanzados, en términos de los principales objetivos que se planea la política o el programa, así como las actividades que lleva a cabo para lograrlos, parece sólo uno de los primeros pasos de la evaluación.

El segundo orden de análisis propuesto por Fischer sería más cercano a los aspectos que destaca la perspectiva de los sistemas complejos pues comienza a explicitar el entramado social de la política o programa en cuestión; es decir, indicaría los posibles impactos de la política en el marco instrumental y societal en que tiene lugar el problema, así como los bienes y servicios encomendados a tratarlo. Por su parte, la dimensión valoral, enfatiza la materialización los valores legítimos de la sociedad que reciben los programas y las políticas. La información producida en estos dos últimos niveles sirve para analizar con mayor profundidad condicionamientos de carácter normativo, institucional y de procesos de las políticas.

En todos los casos, los resultados de la evaluación tienden a utilizarse para apoyar decisiones en las dependencias o instituciones sectoriales y buscan comprender en qué medida están operando en la realidad ciertos valores y hasta dónde se está incidiendo en determinado sistema social, pero que *strictu sensu* son ajenas al diseño y a la operación de la política o programa en cuestión.

Como nivel de evaluación, la contribución social en el sistema, así como el alcance de ciertos objetivos nacionales traducidos en valores, puede detectarse mediante la suma del conjunto de políticas que abonan al mismo ámbito, lo cual requiere de un sistema de información con los mismos estándares para ser capaz de recibir datos de naturaleza diversa y compararlos, además de tener la capacidad de identificar los aportes en la explicación causal de cada una de las acciones políticas involucradas en la resolución del problema. Objetivos de esta naturaleza son difíciles de

plantear en un contexto donde, en la práctica, cada programa o segmento administrativo opera como “islote”, a pesar de la referencia a los objetivos nacionales del programa de desarrollo o de los programas sectoriales.

El fin último de los programas gubernamentales es la resolución de problemas concretos, lo cual exige cercanía con teorías, conceptos o perspectivas normativas, además de herramientas administrativas. Para el cálculo o aprecio del logro de estos objetivos, la política debe producir resultados específicos medidos cualitativa y cuantitativamente, y valorados respecto del sistema social al que pertenece. En este sentido, el primer y segundo órdenes de la evaluación de la política permiten la acumulación de aspectos relevantes de los programas, lo cual da a lugar a redefinir y mejorar sus cursos de acción, sus objetivos e incluso los recursos que ahí intervienen.

La planificación y sistematización de los problemas públicos bajo teorías de sistemas complejos permite prever e identificar en dónde y cómo podemos minimizar la sorpresa; de ahí la importancia de abrir espacios para la recolección de información adecuada y la búsqueda de procesos de rediseño y evaluación más flexibles. De otro modo, el formato y procesos uniformes pueden ser incapaces o, en el mejor de los casos, limitados ante la variedad de contextos y entornos cambiantes.

En este sentido, las soluciones únicas a los problemas y las estructuras uniformes para la evaluación, si bien útiles en algunos sentidos, como los que ya he señalado antes, encuentran sus límites intrínsecos para el diseño en los siguientes aspectos: 1) el contenido del programa en sí mismo; 2) la fidelidad entre el diseño y la realidad en la que se trata de influir; 3) la robustez de la formulación ante la variedad del contexto o entorno; 4) el conocimiento del contexto del programa, y 5) el tiempo entre la operación y los resultados observados (cfr. Morell, 2010).

Adicionalmente, en la etapa de implementación del programa se produce no sólo información sobre la cercanía entre lo planificado con lo que sucede en la realidad, sino que también se ofrece conocimiento sobre los límites para obtener la información que con anticipación se consideró como adecuada, así como la validez de los marcos de interpretación de esos datos. En este sentido, la identificación y caracterización de las siguientes variables pueden ser útiles para dimensionar la sorpresa: número de involucrados en el problema y programa, incluidos los diversos niveles de gobierno; tipo de bienes y servicios que provee el programa, periodos de tiempo para la

valoración de resultados y uso de los resultados de la evaluación. Estos son aspectos que posteriormente podrán guiar los cambios en el diseño, la ejecución y la valoración de resultados o evaluación.

A partir de lo anterior es posible conducir una retroalimentación entre los marcos teóricos-conceptuales previstos con la experiencia. Recordemos que las decisiones, en un futuro, dependen de la interpretación de acontecimientos pasados. Es lo que Morell identifica como conocimiento del proceso, es decir, las condiciones bajo las cuales opera el programa, y el conocimiento del dominio, aprendizaje a partir de programas similares y su aproximación a los resultados y a los hechos del programa que se está llevando a cabo. Entre el planteamiento teórico, el diseño, la implementación y el resultado hay una distancia causal y temporal entre problemas y productos, donde el programa público representa el mediador (Morell, 2010). La distancia causal refiere al número de eventos necesarios entre las actividades (incluyendo los recursos disponibles) y los resultados planteados. La distancia temporal es el lapso para que ocurran esos eventos. En ambos casos, mientras más eventos o mayores plazos de tiempo, mayor será la posibilidad de que se presenten cambios imprevisibles o sorpresas. De ahí que las formulaciones de los programas y los diseños de evaluación sean los medios para minimizar de modo eficiente y efectivo los procesos necesarios para alcanzar un resultado, e identificar los parámetros más adecuados para juzgar lo que sucedió en los hechos y los resultados que se produjeron.

El análisis por niveles de la política nos ha permitido distinguir la complejidad de los problemas a partir de grados en los cuales existen lógicas y dinámicas particulares que son relevantes para comprender los resultados de un programa pero que sólo podrían captarse con la incorporación de preceptos de mayor alcance como los que se plantean en una perspectiva sistémica. Bajo ese orden de análisis sistémico y complejo de la política pública se esperaría que, tanto diseños como modelos de evaluación, no sólo se integraran y transitaran de un nivel a otro como parte de su evolución y mejoramiento, sino que también permitieran comprender las interrelaciones entre los involucrados y sus procesos.

En este sentido, y para el caso mexicano, el sistema de monitoreo y evaluación adoptó los beneficios de la compresión lógica y causal de los problemas que atiende, mediante la utilización de la MML y la matriz de indicadores. Sin embargo, la herramienta se juzga limitada pues no está

hecha para identificar la incertidumbre y las relaciones entre los diversos agentes en contextos cambiantes. Un ejemplo de ello es la distancia, hasta cierto punto previsible por la teoría, atribuible a la diferenciación de quien dirige la política y quien la lleva a cabo. Dicho de otro modo, a nivel local, las dinámicas de ejecución pueden generar sus propios sistemas complejos, lo cual da lugar a un gran entramado de sistemas desiguales y con sus propias particularidades.

De acuerdo con una perspectiva sistémica de la evaluación, el marco lógico se basa en relaciones lineales (insumo-actividad-producto-resultado) donde aparecen sólo los factores de riesgo que son previsibles (en la columna de supuestos); no es posible plasmar procesos emergentes, ni tampoco es sencillo incorporar las lógicas de niveles administrativos distintos, para el caso de políticas que operan en sistemas descentralizados, a menos que recurramos a matrices de indicadores bajo algunas modificaciones.

Uno de los principales aportes que ofrece esta visión es la posibilidad de ir más allá de los niveles técnico-analítico y contextual para reconocer la presencia de agentes diversos cuyo análisis y comprensión de su comportamiento difícilmente se alcanzará como la suma de acciones y actividades estandarizadas debido al ambiente en que se desarrollan, así como el reconocimiento del entorno cambiante en que se sitúan los problemas y la operación de las políticas y programas diseñados que exige a los actores involucrados aprender, retroalimentar y evolucionar en torno a una situación particular.<sup>18</sup>

Con base en lo anterior, podemos señalar que en la evaluación, el orden y la profundidad con que se estudian los problemas y las políticas importan porque a partir de ello pueden vislumbrarse la mejora en los diseños de políticas y programas, así como los retos para evaluar los resultados y los alcances logrados. Pero lo más importante es que permite identificar que la complejidad está presente en la acción gubernamental, así como en sus estrategias, que los sistemas pueden ser difíciles de capturar en modelos lineales simples, y que las previsiones de cambiar y mejorar pueden ser admisibles en el proceso de la política pública y su evaluación. Dicho de otro modo, bien pueden buscarse modelos de evaluación que sean dinámicos, flexibles y que evolucionen junto con los problemas que pretenden valorar.

La búsqueda de metodologías adaptables a los problemas sociales resulta una vía de solución factible para manejar la incertidumbre y seguir

buscando la mejora de los programas sociales con resultados integrales. Si miramos los modelos de marco lógico de las políticas educativas nacionales que presentamos antes, puede concluirse que tanto la complejidad de los sistemas, como los problemas de decisiones de niveles gubernamentales traslapados, como los sucesos no esperados, no están incorporados; y, en el mejor de los casos, se les menciona en la columna de “supuestos”, sin mayor especificación o atención por parte de la acción pública, ni de la evaluación.

Desde el punto de vista de la construcción de información relevante, y trayendo de nuevo la discusión sobre el uso de la matriz de marco lógico en lo que respecta al seguimiento de resultados y la utilización de indicadores, encontramos que hay una tendencia a confirmar objetivos que, muy posiblemente, no han sido legitimados como necesidades o demandas sociales. Lo más recurrente es que las instituciones se preocupen por desarrollar mejores herramientas que midan correctamente los logros previos de los programas, incluso cuando se ambiciona alcanzar una valoración de los efectos con base en la eficiencia, la eficacia y la calidad. Se hace hincapié en las mejores alternativas para evitar la influencia de factores ajenos al programa que puedan afectar la medición; no se realizan los mismos esfuerzos para otorgar significación en las mediciones y resultados a partir de valores sociales y perspectivas sistémicas, como si la búsqueda de resultados no fueran instrumentales a la sociedad en su conjunto, sino sólo a la lógica administrativa y presupuestal.

Por otra parte, la posición política que asumen los actores en una política puede ayudar a comprender, en parte, la tendencia a la búsqueda de resultados sin sustento sistémico y valoral. Recordemos que los resultados de la evaluación tienen un auditorio, en palabras de Majone (1997:79), un conjunto de receptores y usuarios efectivos o potenciales del análisis, o tomador de decisiones, con diferentes fines, posturas e intereses.

En este sentido, la evaluación tiene que ser concebida como un “proceso continuo de aprendizaje recíproco” (Majone, 1997:79), en el que no se trata sólo de recopilar información, analizarla y difundirla a través de un informe con fines administrativos. Se trata de avanzar en el conocimiento, análisis e interpretación hasta llegar a una mejor comprensión de los problemas públicos para conseguir mayores efectos de las acciones públicas en la resolución de las problemáticas sociales. Sólo así es que se podría alcanzar un efecto a nivel empírico y normativo de la política.

De acuerdo con esta idea es imprescindible allegarnos de datos relevantes, con bases normativas, producir mejores soluciones, apoyar la toma de decisiones y reorientar la acción pública. Propiciar la cercanía con beneficiarios porque de ellos se recibe la mejor información sobre la política en cuestión y entonces poner en práctica el conocimiento adquirido. Visto así, la evaluación es el punto de encuentro de diversos intereses, un diálogo entre académicos y políticos, en el que tiene que haber un estrecho vínculo entre las funciones y los usos de los resultados obtenidos de ella (Reimers y McGinn, 1997).

### **¿Alcanza un método para emitir juicios sobre las políticas educativas?**

He centrado la discusión sobre la MML para la evaluación de las políticas educativas y sobre algunos límites que se impusieron en su inserción como herramienta para la evaluación de las acciones gubernamentales en el país. Destaqué también sus virtudes en términos de su simplicidad, comparabilidad y uniformidad de la información, entre otras. Sin embargo, he propuesto la inclusión de una perspectiva sistémica de la evaluación, que sea capaz de mantener flexibilidad y apertura en los modelos de evaluación para alcanzar mayor comprensión en el análisis de los problemas públicos y los diferentes actores que intervienen en ellos. Cada programa público, al menos en el ámbito educativo, permitiría desarrollar modelos de evaluación que fuesen más productivos para el propio desarrollo de los programas y acciones gubernamentales.

¿Qué hemos logrado con el uso del marco lógico? La evaluación, y poco a poco también los diseños o la reformulación de los programas funcionan cada vez más de acuerdo con una estructura causal. La tarea siguiente es ampliar la visión tradicional del monitoreo de los programas, incluyendo los cambios en su lógica, y tener más datos acerca de objetivos y actividades en un entorno diverso y cambiante. Esto con el fin de adaptar las técnicas de comprensión, interpretación y control en los equipos de planificación, implementación y evaluación de los programas.

Los problemas públicos, considerando el entorno cambiante y complejo en que se desarrollan, además de los diversos agentes involucrados y los cambios en las intervenciones (e incluso apropiaciones) en el espacio público, exigen procesos de aprendizaje y adaptación constante de la acción gubernamental, de modo que los métodos de evaluación no podrían conducirse de otra manera.

La principal misión de la investigación evaluativa es producir información que coadyuve a la toma de decisiones y a la mejora de las políticas y los programas bajo marcos que permitan captar la misión y el significado de las acciones públicas de acuerdo con un conjunto de valores congruentes a necesidades y demandas sociales; si este vínculo con los valores que dan motivo a la existencia de la acción de las políticas no se logra, se producen informaciones y resultados *ad nauseam*, sin ningún propósito más que la legitimación, o la crítica fatalista, de los programas gubernamentales, sin alcanzar el objetivo de retroalimentación a la toma de decisiones; no importa cuán sofisticadas sean las metodologías, ni cuán comparables sean los resultados con otros programas de gobierno o, incluso, con el ámbito internacional. Los diferentes discursos de análisis sobre la política y los sistemas complejos coinciden en la importancia de integrar hechos y valores, ya que si alguno de estos elementos está ausente, la interpretación o la medición de la información estará incompleta y su aportación a los propósitos de la evaluación de la política podría estar alejada de ser un “buen juicio”.

El análisis y la evaluación de políticas tienen frente a sí un importante reto: la creación de modelos creativos, flexibles, incluyentes y sistémicos para lograr un acercamiento con lo imprevisible de los contextos y las relaciones entre involucrados, teniendo como herramienta la comprensión de los problemas desde el punto de vista teórico, el diseño de procesos congruentes con ese marco y la identificación del sistema y los valores relacionados. No cabe duda que este marco de análisis conlleva importantes exigencias respecto del tiempo, costo y manejo de la información.

La reflexión en torno a los niveles de evaluación y la adición de una perspectiva valoral exige transitar de un compromiso a otro de la evaluación, lo cual no va a lograrse con una sola herramienta, sino con un conjunto de métodos concurrentes y complementarios. Para lo cual es atinado comenzar con el conocimiento de los límites de metodologías como el marco lógico y la matriz de indicadores, para resarcir las faltas en que se suele incurrir en el proceso de evaluación de las políticas actuales.

Sin embargo, no hay que olvidar que para llevar a la práctica los modelos evaluativos y hacer de sus resultados evidencia para la toma de decisiones, es imprescindible la comunicación y consenso con los actores ajenos al diseño, la implementación y evaluación de los programas para incluir sus demandas, necesidades y su visión de la política que, en suma, representan una forma más para delimitar los valores que está generando la política o programa público.

La responsabilidad del evaluador es realizar su análisis con la mejor información posible, con los marcos teóricos conceptuales más adecuados a la problemática que ocupa a los tomadores de decisiones e identificar, a través de ellos, el conjunto de valores sociales que dan referencia a la intervención evaluada. Consiste también en detectar los alcances y límites de los modelos de evaluación existentes, así como apoyar el uso de la evidencia para la mejora de la acción pública, la realidad social y las condiciones de vida de las personas.

### Notas

<sup>1</sup> Aunque pudiese aplicarse a otros espacios de las políticas sociales de manera general.

<sup>2</sup> Véase, como ejemplo, el modelo de evaluación de la calidad educativa del Instituto Nacional para la Evaluación de la Educación (INEE) en *Panorama Educativo de México*, 2008.

<sup>3</sup> De hecho, la MML y la construcción de su respectiva MIR constituyen sólo una de las variantes del uso de modelos lógicos para determinar las relaciones causales y lógicas así como los efectos de los problemas públicos (véase, entre otros, Frechtling, 2007 y Bamberger, *et al.*, 2006).

<sup>4</sup> Es razonable pensar que tanto el Coneval como la Secretaría de Hacienda y Crédito Público (SHCP) hayan elegido este modelo de evaluación por esta virtud comparativa general.

<sup>5</sup> Tales como el uso de matrices anidadas (o en cascada) o la adición de una quinta columna para el seguimiento de los supuestos que he propuesto antes.

<sup>6</sup> Por ejemplo, la inclusión de un componente ajeno a la lógica causal de un programa, debido a que administrativamente ya estaba operando bajo esa misma instancia administrativa.

<sup>7</sup> Se incluyen programas coordinadores por las Subsecretarías de Educación Básica, Media Superior y Superior, el Instituto Nacional para la Educación de los Adultos (INEA), Consejo Nacional de Fomento Educativo (Conafe), Comisión Nacional de Cultura Física y Deporte (Conade) y Consejo Nacional para la Cultura y las Artes (Conaculta), así como el programa Desarrollo Humano Oportunidades, en el cual la SEP participa de manera directa.

<sup>8</sup> Se tomó en cuenta en el análisis tanto la MIR como los resultados de la evaluación exter-

na publicados en: [http://www.sep.gob.mx/es/sep1/programas\\_federales\\_sujetos\\_a\\_evaluacion\\_externa](http://www.sep.gob.mx/es/sep1/programas_federales_sujetos_a_evaluacion_externa)

<sup>9</sup> A partir del TR1: Modelo de Términos de Referencia para la Evaluación de Consistencia y Resultados, y en su caso el TR2: Modelo de Términos de Referencia para la Evaluación de Diseño, se retomó la información del punto “2.1.1 Características del programa” a fin de obtener la siguiente información: año de inicio de operación, modalidad, dependencia coordinadora del programa y la participación (o no) de las entidades federativas, el Distrito Federal, y los municipios. Del apartado 2.1.2 Análisis de la contribución del programa a los objetivos estratégicos de la dependencia y/o entidad, así como a los objetivos nacionales, se retomó la información de las preguntas 1 y 2 (¿El problema o necesidad prioritaria al que va dirigido el programa está correctamente identificado y claramente definido?, y ¿Existe un diagnóstico actualizado y adecuado, elaborado por el programa, la dependencia o entidad, sobre la problemática detectada que sustente la razón de ser del programa?, respectivamente) para calificar de forma binaria si el programa identificó el problema que atiende y si contaba con un diagnóstico actualizado del mismo.

<sup>10</sup> Para recolectar la información sobre la población objetivo y potencial se retomó la información de los Modelos de Evaluación TR1 y TR2, preguntas 23 y 24 (¿La población que presenta el problema y/o necesidad (población potencial), así como la población objetivo están claramente definidas?, y ¿El programa ha cuantificado y caracterizado ambas poblaciones, según

los atributos que considere pertinentes? (En el caso de individuos, en términos de edad, sexo, nivel socio-económico –señalar quintil de ingreso si corresponde–, principales características de la actividad económica que desempeña –rama de actividad, condición de empleo, etc.–, condición indígena u otros atributos que sean pertinentes, respectivamente), para determinar en qué medida el programa había definido, caracterizado y cuantificado su población objetivo y potencial. De este modo, cuando los programas cubrían ambos aspectos, se calificó como “Sí”; la falta de ambos aspectos se denominó como “No”, y cuando contestó afirmativamente a alguna de las dos preguntas, se determinó como “Medio”.

<sup>11</sup> De los modelos de evaluación antes mencionados, punto 2.1.3 Evaluación y análisis de la matriz de indicadores, se retomó la pregunta 17 (¿El programa ha identificado los medios de verificación para obtener cada uno de los indicadores?) para obtener la fuente de información y la frecuencia de medición, concentrándonos en los niveles de fin y propósito, esto en caso de que no contar con las MIR.

<sup>12</sup> Sólo se cuenta con el dato de frecuencia para 24 de un total de 40 evaluaciones.

<sup>13</sup> Este es el caso de los programas de la Subsecretaría de Educación Básica y el INEA.

<sup>14</sup> Es el caso de las evaluaciones realizadas sobre educación superior y posgrado y de algunos programas de becas. Además de producir inconsistencias al momento de aplicar los Modelos de Evaluación de Diseño y Consistencia y Resultados del Coneval que impedirían un juicio válido de la operación por parte de las instituciones educativas o unidades responsables porque la identificación del problema que atiende el programa, como su diagnóstico, así como la definición, caracterización y cuantificación de las poblaciones objetivo y potencial tendrían que realizarse de manera independiente, no de manera global como lo demandan dichos lineamientos, además de opacar la relación resultado y uso de los recursos

<sup>15</sup> Ante estas limitaciones sería acertado preguntar cuándo y para qué tipo de problemas es adecuada la metodología, qué podemos esperar de su uso y qué vamos a favorecer en el diseño y la evaluación de las políticas y programas: el

aprendizaje, la rendición de cuentas o, incluso, el consenso y la negociación.

<sup>16</sup> Dejo fuera otras muchas discusiones sobre evaluación, incluyendo la que refiere a la definición de los actores relevantes en la política. Valga mencionar, por ejemplo, el trabajo de Subirats *et al.*, 2008, en que se señala que los actores implicados en las políticas públicas disponen de recursos para defender sus posiciones y es el contexto institucional el que influye en sus pautas de comportamiento tanto de manera individual como colectiva. De este modo, el autor propone como división para el análisis la presencia de actores, recursos e instituciones. En este tenor, se considera que un conjunto de individuos constituirá un actor unitario cuando se presente y actúe con homogeneidad interna respecto a los valores e intereses que defiende y los objetivos que persigue. Por ello, hay actores intencionales, pues existe un marco de maniobra mayor o menor en cada situación en donde éstos se desenvuelven. Bajo esta concepción, los factores institucionales o sociales no determinan de manera absoluta y lineal las apreciaciones, decisiones o conductas de los actores públicos o privados (parcialmente) autónomos (Subirats, *et al.*, 2008: 51-54).

<sup>17</sup> En Bracho (2010) se encuentra una ejemplificación de los distintos niveles de evaluación para el caso de la evaluación de los adultos y el tema de equidad.

<sup>18</sup> Morell enfatiza que las sorpresas y lo inesperado responden, principalmente, a eventos que pudieron ser previstos (en mayor o menor medida) si un mecanismo de cambio posible se hubiese previsto en el modelo de la política, o bien, se deben a acontecimientos que, por razones teóricas, pueden preverse como productos emergentes en el tiempo. Lo cierto es que hay que atenuar dichos imprevistos y sorpresas, así como los métodos para lograrlo y, por otro lado, estar ciertos de los límites de nuestras herramientas metodológicas cuando enfrentamos sistemas y comportamientos de tales características. Pues a medida que nos alejamos de lo previsible, menor será la capacidad para identificar relaciones importantes para diseño, el funcionamiento y la evaluación de las políticas, corriendo el riesgo de difuminarse.

## Referencias

- Bamberger, Michael; Jim Rugh y Linda Mabry (2006). *RealWorld evaluation: working under budget, time, data, and political constraints*, Thousand Oaks: Sage Publications, Inc.
- Bracho, Teresa (2009a). *Desafíos de la evaluación por resultados de los programas sociales en México*, Ciudad de México: mimeo.
- Bracho, Teresa (2009b). “Evaluación de políticas educativas. Retos y perspectivas”. *X Congreso Nacional de Investigación Educativa. Conferencias Magistrales*, Ciudad de México: COMIE, pp 467-482.
- Bracho, Teresa (2010). “Niveles de análisis en la evaluación de políticas educativas. Insumos para reflexionar en la EPJA”, *Decisio. Saberes para la acción en la educación de adultos*, número 25.
- Bracho González, Teresa; Jimena Hernández Fernández y José Manuel del Río Zolezzi (2009). “Comentarios al “Modelo de Términos de Referencia para la evaluación en materia de diseño, TdR (versión piloto 1.0) de Coneval”, *Uso de los TdR en programas presupuestales complejos*, Ciudad de México: Coneval.
- Castro Fernando, Manuel; Gladys López-Acevedo; Gita Beker Busjeet; Ximena Fernández Ordoñez (2009). “Mexico’s M&E System: Scaling Up from the Sectoral to the National Level”, *ECD Working Paper Series*, núm. 20. Washington, DC: The World Bank.
- Coneval (2010). “La evaluación de los programas sociales: un proceso de mejora continua y rendición de cuentas”. Disponible en: <http://www.coneval.gob.mx/cmsconeval/rw/pages/evaluacion/procesodeevaluacion.es.do> (Consultado en diciembre de 2010).
- Coneval, SFP y SHCP (2010). TR1: Modelo de Términos de Referencia para la Evaluación de Consistencia y Resultados. Disponible en: [http://www.coneval.gob.mx/cmsconeval/rw/resource/coneval/eval\\_mon/906.pdf](http://www.coneval.gob.mx/cmsconeval/rw/resource/coneval/eval_mon/906.pdf) (Consultado en diciembre de 2010).
- Fischer, Frank (2006). *Evaluating public policy*, Ciudad de México: Thomson/Wadsworth.
- Frechtling, Joy (2007). *Logic modeling methods in program evaluation*, California: John Wiley & Sons.
- Gasper, Des (2000). “Evaluating the ‘logical framework approach’ towards learning-oriented development evaluation”, *Public Administration and Development*, 20, pp. 17-28.
- González Arreola, Alejandro (Coord.) (2008). *¿Gobernar por resultados? Implicaciones de la política de evaluación del desempeño del gobierno mexicano*, Ciudad de México: GESOC, AC.
- INEE (2008). *Panorama Educativo de México*, Ciudad de México: INNE.
- Martínez Rizo, Felipe y Annette Santos (2009). “La evaluación educativa”, en: De Alba, A. y Glazman, R. (coord) *¿Qué dice la Investigación Educativa?*, Ciudad de México: COMIE.
- Morell, Jonathan A. (2010). *Evaluation in the face of uncertainty: Anticipating surprise and responding to the inevitable*, Nueva York: The Guilford Press.
- Majone, Giandomenico (1997). *Evidencia, argumentación y persuasión en la formulación de políticas*, Ciudad de México: Colegio Nacional de Ciencias Políticas y Administración Pública.
- Ortegon, Edgar; Juan Francisco Pacheco y Adriana Prieto (2005). *Metodología de Marco Lógico para la planificación, el seguimiento y evaluación de proyectos y programas*, Santiago de Chile: CEPAL.

- Reimers, Fernando y Noel McGinn (1997) *Informed dialogue: Using research to shape education policy around the world*, Westport, Connecticut: Praeger Publishers.
- Rossi, Peter H.; Mark W. Lipsey; Howard E. Freeman (2004). *Evaluation. A Systematic Approach*, California: Sage Publications, Inc.
- Roth Deubel, André-Noel (2006). *Políticas públicas*, Bogotá: Ediciones Aurora.
- Sánchez Ramírez, Mariana Anabel (2011). *El uso de la Metodología de Marco Lógico en el diseño y la evaluación: alcances y limitaciones. El programa Hábitat y el Plan Municipal de Desarrollo, Puebla 2008- 2011*, Ciudad de México: Flacso-Sede México.
- Secretaría de Hacienda y Crédito Público, Secretaría de la Función Pública y Consejo Nacional de Evaluación de la Política de Desarrollo Social (2007). "Lineamientos generales para la evaluación de los Programas Federales de la Administración Pública Federal", Ciudad de México: *Diario Oficial* de la Federación, 30 de marzo de 2007.
- SEP (2010). *Evaluación externa de programas federales*. Disponible en: [http://www.sep.gob.mx/es/sep1/programas\\_federales\\_sujetos\\_a\\_evaluacion\\_externa](http://www.sep.gob.mx/es/sep1/programas_federales_sujetos_a_evaluacion_externa) (Consultado en diciembre de 2010).
- Subirats, Joan; Peter Knoepfel; Corinne Larrue y Frédéric Varone (2008). *Análisis y gestión de políticas públicas*, Barcelona: Ariel.
- Tiana Ferrer, Alejandro (1997). *Tratamiento y usos de la información en evaluación*, Madrid: Universidad Nacional de Educación a Distancia. Disponible en <http://caedofu.tripod.com/doc/tiana.pdf> (consultado en diciembre de 2010).
- Weiss, Carol H. (1997). *Evaluation*, segunda edición, Nueva York: Prentice Hall.
- Williams, Bob y Richard Hummelbrunner (2009). *Systems concepts in Action. A practitioner's Toolkit*, California: Stanford University Press.

**Artículo recibido:** 6 de febrero de 2011

**Dictaminado:** 27 de marzo de 2011

**Segunda versión:** 9 de mayo de 2011

**Tercera versión:** 2 de junio de 2011

**Aceptado:** 2 de junio de 2011