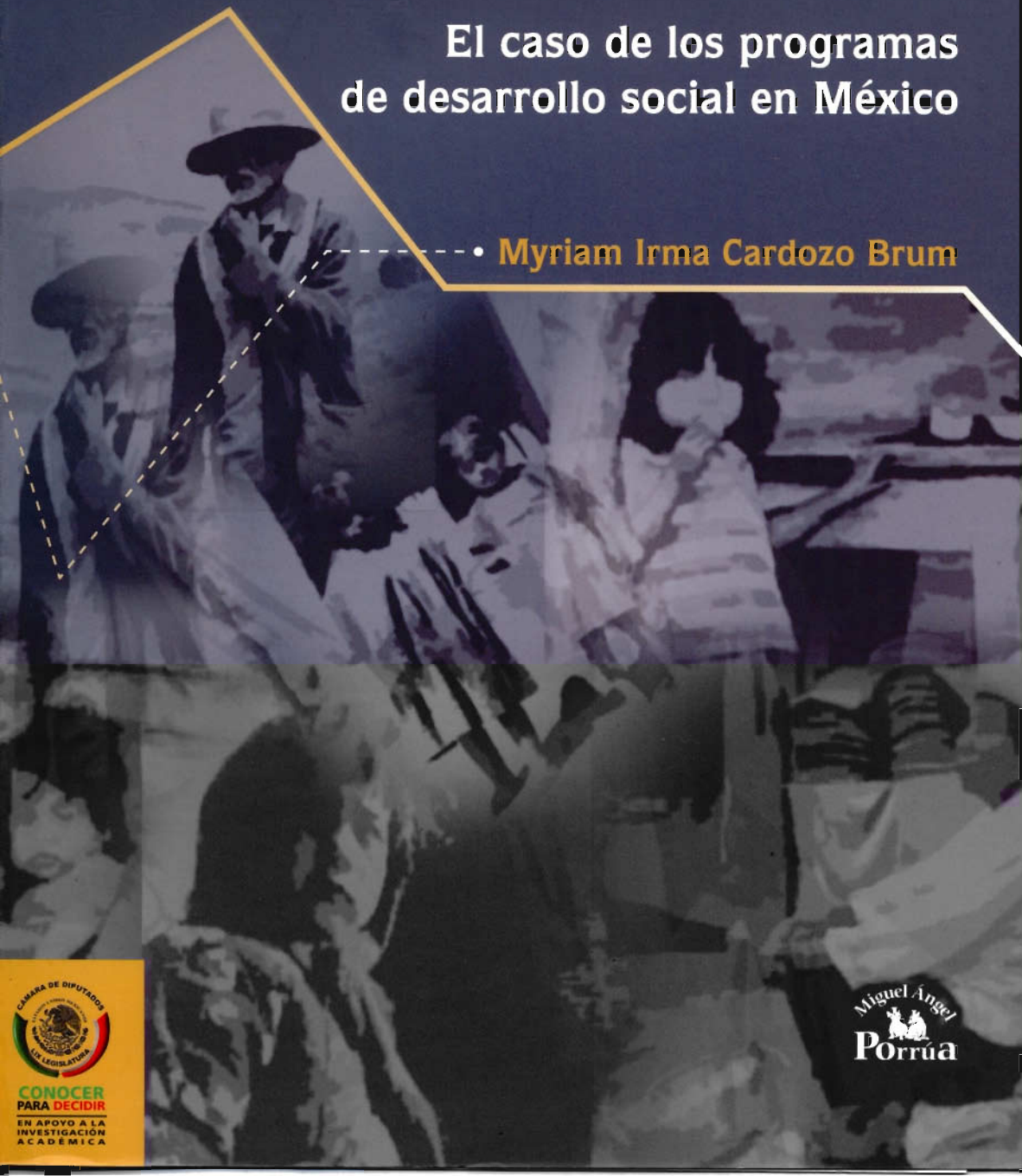


# La evaluación de políticas y programas públicos

El caso de los programas de desarrollo social en México

• Myriam Irma Cardozo Brum

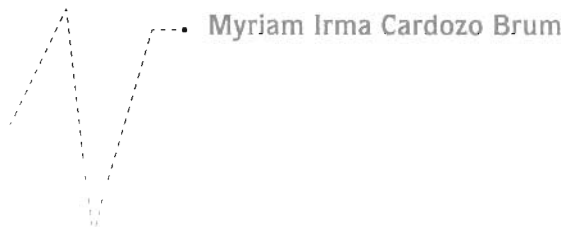


# **L**a evaluación de políticas y programas públicos

El caso de los programas  
de desarrollo social en México

# La evaluación de políticas y programas públicos

El caso de los programas  
de desarrollo social en México



**CONOCER  
PARA DECIDIR**  
EN APOYO A LA  
INVESTIGACIÓN  
ACADÉMICA

Miguel Ángel  
  
**Porrúa**

---

MÉXICO • 2006

La H. CÁMARA DE DIPUTADOS, LIX LEGISLATURA,  
participa en la coedición de esta obra al incorporarla  
a su serie CONOCER PARA DECIDIR

Coeditores de la presente edición

H. CÁMARA DE DIPUTADOS, LIX LEGISLATURA  
MIGUEL ÁNGEL PORRÚA, librero-editor

Primera edición, agosto del año 2006

© 2006

MYRIAM IRMA CARDOZO BRUM

© 2006

Por características tipográficas y de diseño editorial  
MIGUEL ÁNGEL PORRÚA, librero-editor

Derechos reservados conforme a la ley  
ISBN 970-701-824-0

Queda prohibida la reproducción parcial o total, directa o indirecta del  
contenido de la presente obra, sin contar previamente con la autori-  
zación por escrito de los editores, en términos de la Ley Federal  
del Derecho de Autor y, en su caso, de los tratados internacionales  
aplicables.

IMPRESO EN MÉXICO



PRINTED IN MEXICO

[www.maporrua.com.mx](http://www.maporrua.com.mx)

Amargura 4. San Ángel, Álvaro Obregón, 01000 México, D.F.

*A mis hijos Yamandú, Anabí y Laura,  
por su estímulo, apoyo y comprensión*

# Presentación

Luis F. Aguilar

GOBERNAR UNA sociedad implica un conjunto de actividades normativas, políticas y técnicas, cuyo propósito consiste en establecer los objetivos sociales de beneficio general, seleccionar y prescribir las acciones que se consideran idóneas para producir los objetivos deseados, asignar los recursos requeridos para que las acciones prescritas puedan llevarse a cabo, crear sistemas de información que permitan dar seguimiento a las acciones que el personal público y las organizaciones sociales realizan a fin de producir los objetivos, definir un sistema de medición que haga posible conocer con precisión la brecha que pudiera existir entre las acciones practicadas y las prescritas, entre la calidad de los servicios públicos prestados y los deseados o exigidos, entre los resultados sociales esperados y los concretamente alcanzados.

En este conjunto de múltiples operaciones directivas que integran el acto de gobernar o administrar destaca *la evaluación* porque ofrece información clave acerca del grado de desempeño real del gobierno (actividades, productos y resultados) y nos ayuda de esta manera a conocer si la actividad de un gobierno es de calidad, eficaz, eficiente y responsable, si produce resultados de valor para la sociedad y, a la postre, si por ello el gobierno ha sostenido o perdido su legitimidad social. Gobiernos ineficientes y de baja calidad en la provisión de los bienes y servicios públicos, aun si elegidos, representativos y legales, terminan por frustrar expectativas sociales y hacen perder la confianza ciudadana.

En la disciplina de política pública y en la misma práctica gubernamental la evaluación ha sido tradicionalmente una operación informativa y analítica que tiene lugar al final del llamado “ciclo de la política pública” o cada vez que el gobierno practica cortes en la gestión de sus programas. Al colocarse al cierre del ciclo de la política o del programa operativo (anual o multianual) y al señalar en cuánto los resultados producidos por la acción pública se corresponden con los objetivos sociales esperados o exigidos, la evaluación ha sido el mecanismo gubernamental y ciudadano apropiado

para conocer la idoneidad y eficacia de las acciones que el gobierno o el gobierno en unión con la sociedad han emprendido para resolver los problemas presentes de la vida asociada y realizar los futuros deseados de convivencia. Por consiguiente, la evaluación de la acción del gobierno ha representado en algunas condiciones políticas y administrativas un genuino proceso de aprendizaje público que ha ayudado a que gobierno y sociedad identifiquen y corrijan errores y fallos en las políticas (en los objetivos, en los instrumentos, en los tiempos) y a que se decidan a cancelar organismos y programas que han mostrado por años ser ineficientes y hasta contraproducentes en sus resultados. Pero en otras condiciones la evaluación ha sido simplemente un ritual burocrático, que se exige como requisito para armar el nuevo presupuesto gubernamental, o un ejercicio tras bambalinas de los altos ejecutivos del gobierno que les hace conocer sus decisiones equivocadas y sus defectos operativos, pero resultados que no hacen públicos para salvar la cara y tratar de mostrarse como líderes inteligentes y eficientes, a pesar de los magros logros de sus políticas y la mala calidad de los servicios públicos.

Debido a que la evaluación arroja información sobre el grado y alcance del desempeño gubernamental en determinados campos públicos (seguridad pública, salud, educación, infraestructura, empleo, pobreza...), no ha gozado de la simpatía de los gobiernos, que frecuentemente han preferido no practicarla, o la han realizado con métodos de medición laxos o incompletos, o han ocultado al público ciudadano los resultados adversos de las evaluaciones institucionales, financieras, operativas, de calidad del servicio, de impacto social... A pesar de esas eventuales distorsiones y resistencias gubernamentales, la evaluación es el acto directivo por excelencia por cuanto su información empírica objetiva ofrece elementos para conocer la eficacia de las acciones gubernamentales y, en consecuencia, identificar equivocaciones, corregir el rumbo y ajustar metas, instrumentos, tiempos, recursos. Permite asimismo determinar la eficiencia de las acciones de gobierno y ayuda entonces a conocer la relación entre costos y beneficios de su acción y de esta manera está en aptitud de prevenir desperdicios, incentivar ahorros y cancelar operaciones innecesarias.

La novedad actual de la evaluación consiste en que ha dejado de ser un acto directivo más, aun sí muy importante, y se ha convertido en el eje de la dirección de una organización pública. Si en el pasado se gobernaba a una sociedad o se dirigía a una agencia administrativa de gobierno mediante dispositivos de mando y control, hoy los gobiernos actúan en condiciones menos discrecionales y más restringidas, en las que su autoridad está más dispersa, descentralizada, existe mayor profesionalización de los funcionarios

públicos y crece la exigencia ciudadana por la eficacia, la transparencia y la rendición de cuentas. En estas condiciones en las que la autonomía decisoria del mando gubernamental se ha visto reducida y en las que es prácticamente imposible un control cerrado de los mandos y operadores de la administración, la actividad de gobernar o dirigir no puede llevarse a cabo más que mediante información y evaluación de las acciones que realizan las unidades y personas encargadas de prestar determinados servicios públicos o de realizar determinados objetivos públicos. El gobierno puede dirigir sólo a condición de que esté oportunamente informado acerca de las acciones y los resultados que han conseguido los agentes públicos o los sociales en el desarrollo de específicos programas y, además, a condición de que haya evaluado su rendimiento, midiendo la conformidad o disconformidad de los resultados producidos respecto de los resultados socialmente exigidos, su cercanía o lejanía. Sólo con datos confiables de éxito o fracaso operativo, de costos y beneficios, de satisfacción de usuarios, de disminución o aumento de situaciones sociales calamitosas... el gobierno está en condiciones de gobernar, mostrando persuasivamente a unidades descentralizadas y a profesionales políticamente independientes lo bien o lo mal que están haciendo su trabajo público, lo barato o costoso de sus operaciones, la baja o alta correspondencia de sus productos (bienes y servicios) con las expectativas de la sociedad civil, económica o política. Y al hacerlo dispone de las evidencias y los argumentos para defender sus decisiones, así como para asesorar y ayudar, corregir y exigir modificaciones, y también para despedir, cancelar, suspender...

No debe resultar extraño entonces que los gobiernos llamados autoritarios, que se caracterizan por no estar sometidos al escrutinio de una sociedad política y civil independiente, no hayan practicado sistemáticamente la evaluación o hayan tenido que hacerla sólo en momentos de urgencia política, cuando estaban al borde de críticas y movilizaciones sociales severas, o bajo la exigencia estricta de los organismos internacionales financieros a los que recurrían por préstamos con el fin de evitar (posponer) la crisis fiscal del Estado o para superarla después de la caída. Sobran también los ejemplos de gobiernos públicamente responsables o políticamente oportunistas que, deseosos de mejorar su rendimiento frente a la demanda social, procedieron a evaluar de manera permanente o periódicamente sus acciones y resultados. Asimismo, no debe resultar extraño que los gobiernos planificadores rectores de la economía, los que buscaban o pretendían gobernar mediante planes, planes nacionales, estatales o municipales de desarrollo, nunca hayan ligado la planeación con la evaluación para medir su pertinencia por lo menos.



Hemos vivido la experiencia de gobiernos que frecuentemente se han mostrado mejores planificadores que realizadores, expertos en elaborar hermosos planes integrales más que en lograr resultados.

Siempre será posible señalar un antecedente de la evaluación aquí y allá en la historia gubernamental de los países. La verdad es que sólo muy recientemente, en los años ochenta y noventa, después de las crisis financieras de nuestros estados desarrolladores, que desplomaron la economía pública y la economía nacional, apareció la evaluación de la acción del gobierno, un poco a la manera de tratar de cerrar el pozo después del niño ahogado. A pesar de resistencias y simulaciones, la evaluación del desempeño gubernamental, la averiguación de cuánto logran las unidades administrativas y los servidores públicos con los recursos públicos que reciben, se fue abriendo paso en los países como una exigencia directiva inaplazable. La crisis como fecha de su nacimiento determinó también que la evaluación se realizara desde un enfoque predominantemente financiero, interesado en valorar la eficiencia económica de los programas y del personal a fin de empezar a disminuir el gasto público desproporcionado e innecesario y comenzar a restablecer el equilibrio financiero desfondado. El resultado ha sido que la evaluación funcionó como el mecanismo administrativo de las políticas de ajuste y ha sido clave para sanear las finanzas públicas. Otro efecto adicional fue haber afianzado la orientación a resultados (observables, mensurables) como principio de la administración pública.

En este ambiente de reordenamiento de las finanzas públicas fue lógico que la evaluación surgiera en el ámbito de la política social (y especialmente en educación y salud pública), que consume ingentes recursos públicos, y que con frecuencia ha sido utilizada políticamente para ganarse el apoyo popular, sin que el gran gasto público haya podido acreditar haber transformado favorablemente las condiciones sociales de sus destinatarios y haber modificado las situaciones sociales generales más nocivas e inaceptables. Por ello ha sido también lógico que las políticas y los programas sociales sean hoy los que se han visto sometidos obligatoriamente por ley a evaluaciones permanentes que son realizadas por organizaciones civiles y académicas independientes.

El hecho es que en los gobiernos democráticos recientes ya no se gobierna sin que el gobierno evalúe de manera sistemática su desempeño a fin de transparentar el gasto público y rendir cuentas acerca de la legalidad, calidad y efectividad de sus acciones, y sin que utilice los resultados de la evaluación como un proceso de aprendizaje continuo que hace posible corregir errores, defectos, y restablecer la dirección.

De este nuevo proceso directivo, en el que la evaluación ocupa un lugar central, nos habla Myriam Cardozo, apreciada colega que en su trayectoria académica y profesional se ha distinguido porque desde años atrás se ha dedicado a la evaluación de programas, particularmente en el campo de la política social, en un momento donde hablar de la evaluación era más una asignatura académica que una práctica de gobierno. En su libro nos lleva a mirar lo que ha sido y se ha vuelto ahora la evaluación en América Latina y México, que es mirada también a contraluz de las experiencias y prácticas evaluativas de los organismos internacionales y de algunos países desarrollados. Más a fondo, nos lleva a conocer y ponderar los conceptos teóricos fundamentales de la evaluación, sus diversos tipos y usos, sus métodos básicos y técnicas, sus prácticas principales. Aborda cuestiones que están aún a la espera de una respuesta más satisfactoria, como la relación entre los métodos cuantitativos (que suelen ser los más socorridos) y los métodos cualitativos al evaluar, particularmente cuando se trata de política social o de servicios públicos.

Una de las contribuciones más relevantes y pioneras del libro de Myriam Cardozo es su reflexión o escrutinio analítico de la manera como es entendida y practicada la evaluación, por qué y cómo se seleccionan ciertas materias de evaluación y se descartan otras, se seleccionan determinados objetivos en lugar de otros acaso más pertinentes, se emplean ciertos instrumentos y procedimientos de medición y valoración, se derivan determinadas conclusiones. Esta evaluación de la llamada “metaevaluación”, tiene como propósito analizar el proceso de evaluación que practica un gobierno, sus insumos, procedimientos, métodos, resultados e impactos. No tanto o no sólo para hacer la previsible crítica y celebrar que nada funciona en el gobierno, sino para tener una mejor comprensión del trabajo evaluativo, perfeccionar sus operaciones analíticas y procedimientos técnicos, proponer alternativas de mejora. Toda esta serie de metas intermedias persigue un objetivo final que es el consolidar la práctica de la evaluación en las democracias, establecerla como una condición indispensable del gobernar de un gobierno democrático, a fin de evitar que en el futuro se proceda a atenuar o revertir la evaluación, que suele ser el mensajero que trae malas noticias al poder y está condenado por ello a morir.

La primera parte del libro de Cardozo, sustancialmente conceptual, se complementa y expande con la segunda parte, que describe, por un lado, la evaluación de las políticas y programas sociales de desarrollo social en México y, por el otro, lleva a cabo su metaevaluación, revisando los componentes que integran la evaluación social, su enfoque, sus métodos practicados, su organización, sus recursos, sus impactos. Particular atención se presta a los procedi-

mientos que siguen los organismos evaluadores externos de la política social, cuya valoración sólo por el hecho de ser externa o independiente suele considerarse correcta. La mirada se concentra en los antecedentes, las prácticas y el marco regulatorio de la evaluación de los programas de desarrollo social relativamente exitosos en México. Myriam Cardozo construye por sí misma su instrumental metaevaluatorio y al hacerlo nos ofrece un método de valía que puede ser utilizado por nosotros. Su método metaevaluatorio toma su forma más precisa en una lista de comprobación que examina “las decisiones y procesos previos y posteriores a la evaluación de los programas sociales” y que examina además “la metodología y procesos internos de la evaluación”, ofreciéndonos un conjunto de materias e indicadores. La mirada final de Myriam Cardozo es diagnóstica, practica un análisis FODA del sistema de evaluación, y nos señala las perspectivas de la evaluación de políticas (sociales).

Al cerrar el libro de la profesora Myriam Cardozo, puedo decir que he puesto en mi mente orden conceptual, articulando mejor el conjunto de componentes que implica toda evaluación y definiéndolos con mayor precisión. Asimismo, he descubierto la importancia crucial de la metaevaluación. Si la evaluación es el corazón de la actividad de los gobiernos, representa una bienvenida llamada de atención la afirmación de que la evaluación va también evaluada. Sin reflexionar con una mirada acuciosa crítica sobre los supuestos, métodos, formas organizativas, instrumentos y recursos de las prácticas que gobierno y sociedad siguen al evaluar sus programas y servicios públicos, evaluatorias, es probable que los gobiernos caigan en la autocomplacencia gubernamental, erijan dogmas técnicos sobre la evaluación correcta o disfracen con un arsenal cuantitativo y conceptual valoraciones de sus programas hechas a modo.

Como en tantos otros casos, la disciplina de política pública o la de administración pública contribuyen profesionalmente a que los gobiernos y la sociedad desarrollen prácticas razonablemente confiables en términos de método. Pero las dos disciplinas se culminan como disciplinas en sentido estricto cuando retoman su dimensión teórica y reflexiva sobre las prácticas profesionales, las revisan, renuevan y modifican a raíz de los resultados que se desprendieron de un análisis crítico riguroso de las mismas y no de un prejuicio o autoengaño. El trabajo de Myriam Cardozo se mueve en este doble plano, contribuye a la práctica profesional de la evaluación (de la política social) y contribuye a la profundización y enderezamiento de las prácticas profesionales mediante revisión crítica.

## Introducción general

SI BIEN en algunos países industrializados la evaluación de servicios públicos cuenta con un siglo o más de experiencia (Reino Unido, Estados Unidos de América), en la mayoría de los países occidentales la evaluación en el campo de las políticas y los programas públicos ha cobrado importancia en las últimas décadas.

América Latina se ha incorporado recientemente a este proceso con algunos pocos casos nacionales que muestran avances importantes, aunque aún insuficientes. En México, las preocupaciones en torno a la evaluación datan de 1917 y se intensifican a finales de los sesenta; sin embargo, el proceso de evaluación de programas gubernamentales comenzó a instrumentarse de manera sistemática a partir de 1997 con la creación del Programa de Educación, Salud y Alimentación y se generalizó en forma obligatoria para todos los programas gubernamentales sujetos a reglas de operación, con base en la decisión incluida en el Presupuesto de Egresos de la Federación, aprobado por la Cámara de Diputados en el año 2000.

Dicho proceso de evaluación consiste en la realización de una investigación aplicada, de tendencia interdisciplinaria, cuyo objetivo es conocer, explicar y valorar el nivel de logros alcanzado (resultados e impactos) por las políticas y programas públicos, así como aportar elementos al proceso de toma de decisiones para mejorar los efectos de la actividad evaluada.

La necesidad de los gobiernos de aumentar su eficacia en la solución de problemas públicos, hacer un uso más eficiente y transparente de los recursos disponibles y contribuir a la democratización de su gestión, son algunas de las principales razones que explican la relevancia práctica que ha adquirido el tema.

Así, el contar con una evaluación integral permite a los gobiernos direccionar y ajustar estratégicamente sus políticas y programas hacia el logro de

\*Segunda revisión de la tesis "La evaluación de políticas y programas sociales en México: 1998-2004", doctorado en ciencias políticas y sociales con orientación en administración pública, Facultad de Ciencias Políticas y Sociales, UNAM, 2005.

sus objetivos y metas, a partir de la comprensión de sus resultados o *outputs* (productos o servicios) e impactos o *outcomes* (efectos provocados por los productos o servicios en el problema que se trata de resolver). Para ello resulta necesario conocer las características de los recursos empleados (cantidad, calidad, oportunidad, etcétera) y de los procesos que han permitido su transformación en resultados (productividad, eficiencia, etcétera), así como sus principales condicionantes contextuales (económicas, sociales y políticas).

Múltiples autores han señalado la importancia acrecentada que el proceso de evaluación adquiere en el contexto de los países que, en el marco de su disciplina fiscal, disponen de presupuestos particularmente restringidos de recursos para la atención de los problemas vinculados a la pobreza y el desarrollo social. Esto debido al riesgo de no emplearlos en forma totalmente satisfactoria por falta de información que retroalimente el proceso de gestión.

Otro factor que aumenta su importancia lo constituyen los procesos de reforma del Estado recientemente vividos, que incluyen la descentralización de funciones para garantizar la eficiencia y efectividad de su gestión y mostrar una equitativa distribución de los esfuerzos.

La evaluación también resulta útil en procesos de democratización caracterizados por el arribo a los cargos de gobierno de políticos provenientes de diversos partidos, con orientaciones y propuestas diferentes, que necesitan asegurar la congruencia de sus acciones. Dicha congruencia puede resultar aún más difícil de lograr cuando se aspira a desarrollar una gestión incluyente de otros actores sociales (organizaciones de la sociedad civil y ciudadanía en general) como la manifestada en los planes y programas gubernamentales de México.

Finalmente, en países de gran tamaño, la evaluación resulta particularmente relevante para conocer la forma en que se aplican los programas nacionales en los lugares más distantes y de difícil acceso, aun los que cuentan sólo con cierto nivel de desconcentración administrativa, como es el caso de los que se encuentran a cargo de la Secretaría de Desarrollo Social en México.

A pesar de las diversas justificaciones prácticas mencionadas y de la importancia teórica que la evaluación asume en el marco del estudio de las políticas públicas y del New Public Management; en la práctica, a menudo se prescinde de realizarla por múltiples razones (falta de recursos, de tiempo, de consenso sobre sus métodos, su utilidad, etcétera), sus resultados no se difunden o son poco utilizados en la toma de decisiones.

En particular, México cuenta con poca información publicada hasta el presente sobre los procesos de evaluación emprendidos, especialmente los métodos y técnicas aplicados, la relación costos/calidad de los resultados alcanzados, y el nivel de difusión y utilización de dichos resultados en los procesos decisorios.

Ante esta situación, el libro que se presenta ha pretendido, primeramente, poner a disposición de los lectores una síntesis de los conocimientos teóricos y empíricos disponibles en la materia (conceptos, enfoques, métodos y experiencias), y en segundo lugar, aplicar los anteriores para realizar un diagnóstico de la situación que guarda la evaluación de las políticas y programas de desarrollo social en México (metaevaluación) que permita proponer alternativas que coadyuven a su consolidación.

Diversos estudiosos de este tipo particular de evaluación, denominada metaevaluación, plantean que su objetivo es enjuiciar la calidad y el mérito de lo que hacen los evaluadores profesionales, que persigue el perfeccionamiento de sus trabajos y promueve una mejor comprensión de los mismos. Esto implica que la metaevaluación toma como objeto de análisis al proceso de evaluación aplicado a los programas sociales: sus insumos, procedimientos, métodos, resultados e impactos.

En el trabajo que se expone, el cumplimiento del objetivo anterior exigió también incluir y valorar el sistema en que se organizan y desarrollan las actividades realizadas por los evaluadores, incluyendo el papel jugado por otros actores o instancias que también inciden en la calidad del proceso de evaluación desarrollado y del informe final logrado. Se trata de factores de su contexto como: condiciones políticas, legales, culturales, y sobre todo, la organización y operación de las instancias de la Cámara de Diputados y de la administración pública que participan en el proceso.

De esta forma, definimos el trabajo realizado como una metaevaluación que analiza y valora la corrección de los métodos de evaluación, sus técnicas operativas y la calidad de los resultados alcanzados, pero también un conjunto de factores contextuales que afectan a los anteriores, con énfasis en la organización y funcionamiento práctico del sistema que conforma el contexto inmediato de la evaluación.

Como puede verse, la realización de una metaevaluación implica un trabajo complejo debido al alto número de variables implicadas y de relaciones entre ellas, por lo que requiere de diseños metodológicos *ad hoc* en cada caso; sin embargo, su utilidad radica en la comprensión del conjunto, por lo que no tiene sentido profundizar en el estudio especializado de un solo aspecto o un solo programa, ya que esto reduciría sus posibilidades de contribución teórica.

Puesto que al iniciar el trabajo no se contaba con ninguna experiencia de metaevaluación previa que explicitara los métodos utilizados en el tratamiento de la información recogida, fue necesario construir una propuesta metodológica-técnica. La misma combina los métodos disponibles en ciencias sociales, así como las técnicas cuantitativas y cualitativas.

Las técnicas de recolección y análisis de datos utilizadas incluyeron: la revisión documental para reconstruir el estado del arte, los antecedentes y el marco legal en materia de evaluación de políticas y programas públicos; y el análisis de informes de evaluación complementado con la aplicación en campo de 25 entrevistas a informantes-clave para la realización de la metaevaluación. El trabajo de campo realizado permitió ahondar en las percepciones, opiniones, perspectivas y propuestas de los diversos actores relacionados a la evaluación y contribuyó a la búsqueda de explicaciones de los problemas detectados.

Las entrevistas cubrieron un amplio conjunto de actores vinculados a la realización de evaluaciones: autoridades de Sedesol, funcionarios de las instancias operadoras de los programas, funcionarios de otras secretarías e institutos vinculados a la evaluación y rendición de cuentas, evaluadores externos, congresistas e integrantes de organizaciones de la sociedad civil.

Las anteriores permitieron profundizar y complementar la información contenida en los informes, fundamentalmente en relación con los procedimientos administrativos relacionados al proceso de evaluación; y conocer las opiniones de los actores entrevistados sobre los principales problemas enfrentados, los factores que los explican y sus posibles soluciones.

Para el análisis de los informes se construyó una lista de comprobación y criterios de valoración propios a partir de la revisión de diversas propuestas elaboradas por autores e instituciones reconocidas en materia de metaevaluación.

Dicha lista quedó integrada por:

- Decisiones y procesos del contexto inmediato:

*Previos a la evaluación de programas:*

- planeación de procesos de evaluación,
- origen de la decisión de evaluar,
- instancias intervinientes y su papel formal,
- funcionamiento real del sistema:
  - relaciones entre las instancias,
  - licitaciones y términos de referencia,
  - mercado de la evaluación,
  - cumplimiento de obligaciones legales,
  - equipos evaluadores y sus móviles,
  - principales problemas administrativos,
- recursos para evaluar.

*Posteriores a la evaluación de programas:*

- difusión de los resultados de las evaluaciones, y
- utilización de resultados en la toma de decisiones.

- Metodología y procesos internos a la evaluación de programas:

- fuentes de información y análisis de datos,
- pertinencia y diseño del programa evaluado,
- recursos utilizados por el mismo,
- análisis de sus procesos,
- medición de cobertura,
- resultados alcanzados,
- impactos provocados,
- participación y satisfacción social,
- utilización de resultados y de recomendaciones previas,
- valoración global del programa y nuevas recomendaciones,
- escenarios prospectivos,
- informes finales, y
- relación calidad/costo de la evaluación.

Todos estos indicadores fueron valorados con base en criterios generales de investigación (validez, confiabilidad, representatividad, consistencia, objetividad, etcétera) y específicos de la metaevaluación (pertinencia de la evaluación, autonomía del evaluador, legitimidad y credibilidad del proceso, calidad y utilidad de los resultados).

La lista y criterios anteriores se utilizaron en el análisis de cada una de las evaluaciones externas, lo que permitió un posterior estudio del conjunto.

La calidad de las evaluaciones fue medida a través de la elaboración de un indicador compuesto de los tres elementos considerados más importantes para nuestro entorno actual: rigurosidad en el diseño, recolección y tratamiento de datos; cumplimiento de los requerimientos incluidos en los términos de referencia; y explicitación de un juicio valorativo bien fundamentado a la gestión del programa.

El documento que se presenta está organizado en dos partes: *a)* el estado del arte en materia de evaluación de políticas y programas públicos, y, *b)* la metaevaluación de las evaluaciones realizadas a los programas de desarrollo social.

El contenido sintético de los capítulos incluidos en la primera parte es el siguiente:



El primero ubica el papel jugado por la evaluación en el devenir de la administración pública y los enfoques que la han permeado: burocracia tradicional, estudio de las políticas públicas, New Public Management, aportes sociológicos y debates recientes.

El capítulo siguiente sintetiza los principales desarrollos teóricos en materia de evaluación: conceptos, clasificaciones, enfoques e importancia del contexto político-organizativo.

El tercero es un capítulo de carácter metodológico; discute la pertinencia que para la evaluación tienen los diferentes métodos desarrollados en ciencias sociales, sus técnicas y criterios específicos. Analiza la utilidad de los métodos experimentales, cuasiexperimentales y no experimentales, así como los de carácter cuantitativo y cualitativo.

A continuación se dedican dos capítulos a relevar algunas experiencias nacionales e internacionales en la materia: uno de ellos recoge la situación en organismos multilaterales y países desarrollados, y el otro, se refiere a América Latina y México en particular.

Finalmente, los limitados desarrollos teóricos y empíricos disponibles sobre metaevaluación, así como una propuesta para aplicarla en el contexto mexicano, constituyen el objeto del capítulo 6.

La segunda parte corresponde a la realización de la metaevaluación de las evaluaciones aplicadas a los programas de desarrollo social en México, y contiene:

El séptimo capítulo, en que se realiza una indispensable descripción de diversos aspectos que coadyuvarán a la mejor comprensión de la metaevaluación incluida en los siguientes: la evolución de la política social mexicana, el actual Plan Nacional de Desarrollo y Programa Nacional de Desarrollo Social y las estrategias de política de desarrollo social y los programas que la componen.

Los primeros esfuerzos sistemáticos de evaluación en el sector son encarados en el capítulo siguiente. El mismo se refiere fundamentalmente a las experiencias de los programas Pronasol y Progresá en los años 1993-2001.

Las referencias constitucionales, legales y reglamentarias que componen el marco jurídico aplicable a la evaluación externa de programas públicos a partir de 2001, así como las bases y términos de referencia en el caso de los programas de desarrollo social, son tratados en el capítulo 9.

El capítulo siguiente da cuenta de la forma en que se organiza y gestiona el sistema de evaluación externa, resaltando su importancia en tanto contexto en el que se desarrolla el trabajo específico de los evaluadores.

El onceavo es el capítulo central de esta segunda parte, ya que en él se procede a metaevaluar las evaluaciones realizadas a los programas de Sedesol en su gestión 2002, para lo cual se revisan sus métodos, procesos y resultados, con base en los indicadores incluidos en la lista de comprobación propuesta en el capítulo 6. Lo anterior se complementa con una aproximación al estudio de la evolución de algunas de ellas en 1998-2003.

Se sintetizan las conclusiones, identificando los impactos y perspectivas de la evaluación de programas de desarrollo social en el capítulo 12. Para ello se ordenan y analizan las relaciones entre los principales problemas, se proponen soluciones a los mismos y se diseñan escenarios prospectivos bajo distintos supuestos.

Entre las conclusiones destacan:

- El importante desarrollo alcanzado en los últimos años en materia de evaluación de políticas y programas públicos en México, mismo que aún presenta problemas relacionados con la organización y operación del sistema de evaluación, con la selección y aplicación de métodos de evaluación, y la implementación de mecanismos de participación social.
- Una multiplicidad de factores que contribuyen a explicar esta situación, entre los que figuran:
  - El gobierno destina presupuestos muy desiguales para la realización de evaluaciones, que no siempre se correlacionan con los alcances de los programas respectivos.
  - La organización actual del sistema de evaluación no garantiza la independencia de los evaluadores externos, la plena transparencia de los procesos y la rendición de cuentas a la ciudadanía.
  - Muchos de los evaluadores externos, funcionarios de programas y congresistas cuentan con una formación profesional y una experiencia poco desarrollada en el campo, lo que también limita la calidad de sus trabajos de evaluación realizados y la utilización de sus resultados.
  - Los principales documentos emitidos por los poderes Ejecutivo y Legislativo en la materia enfatizan la importancia de la participación social y ciudadana en los procesos de gestión y evaluación de programas, pero los mecanismos diseñados son insuficientes para garantizar su operacionalización.

Las propuestas para coadyuvar a la resolución de los problemas que se presentan han sido planteadas con la intención de que México aproveche la

enorme oportunidad que tiene actualmente de mejorar la calidad de los trabajos de evaluación e integrar sus resultados a la gestión pública, como forma de incidir más efectivamente en el alivio del más urgente problema nacional: la pobreza en que vive la mitad de sus ciudadanos.

En la medida en que las metaevaluaciones se encuentran aún muy poco difundidas y no cuentan con métodos específicos desarrollados, se espera haber contribuido a demostrar su utilidad en tanto proceso social de aprendizaje y a proponer un primer método de trabajo para nuestro entorno, mismo que otros esfuerzos irán perfeccionando con el tiempo.

Primera parte

EVALUACIÓN DE POLÍTICAS  
Y PROGRAMAS PÚBLICOS

# Capítulo 1

## Papel de la evaluación en el contexto de la administración pública

### Introducción

DURANTE el siglo XX la evaluación fue teóricamente reconocida como una de las funciones fundamentales que integran los procesos de la administración pública; sin embargo, fue recién en los ochenta cuando su necesidad práctica parece haberse impuesto en la mayoría de los países.

Para comprender su evolución será necesario revisar los distintos enfoques que han caracterizado a la administración pública, partiendo de la burocracia tradicional, resaltando la importancia que adquirió en el contexto de las políticas públicas y con el advenimiento del New Public Management, así como las aportaciones realizadas por la escuela sociológica y los debates actuales.

Este capítulo pondrá especial atención a la evaluación de resultados de la aplicación de políticas y programas públicos.

### La administración pública tradicional

La administración pública, como práctica, puede ser identificada desde el desarrollo de las primeras civilizaciones y se desarrolla con los grandes imperios. Ya en la época moderna, Woodrow Wilson (1887) la definió como el gobierno en acción, o sea, en términos de R. Uvalle,<sup>1</sup> el conjunto de decisiones y acciones complejas, orientadas a la obtención de resultados, que concretan la capacidad de gobernar para la persecución manifiesta del bienestar social y la implícita de protección al gobierno mismo.

Esta cara visible del gobierno, asumió rasgos de tipo burocrático durante la mayor parte del siglo XX, a menudo sensiblemente diferentes del tipo ideal weberiano (especialmente en América Latina). La administración pública burocrática se caracterizó por una clara distinción entre medios de

<sup>1</sup>Notas de la conferencia "Institucionalidad y Administración Pública", presentada por el doctor Ricardo Uvalle Berrones, dentro del seminario obligatorio de doctorantes de la Facultad de Ciencias Políticas y Sociales, el 27 de mayo de 2003.

administración públicos y privados, los primeros de los cuales se dirigían a la satisfacción de las necesidades colectivas, vía servicios públicos. Se la describe como basada en: formación profesional, criterios objetivos, división del trabajo, estructuras jerárquicas, apego a leyes y reglamentos detallados, especificación minuciosa de funciones y actividades, autoridad y lealtad, carrera administrativa, desempeño neutral, impersonal, autorización detallada de gastos, orientado al servicio público, con control central y directo y promoción como único incentivo.

En la segunda mitad del siglo pasado, la burocracia fue generalmente identificada con sus rasgos más negativos. De esta forma, la administración pública tradicional fue acusada de: excesivo crecimiento, centralismo, subutilización de su estructura formal, poca coordinación, incapacidad para atender las demandas, ingobernabilidad, lentitud, ritualismo y rigidez en los procedimientos, apego irrestricto a la norma, la rutina y el formalismo, exceso de control, insuficiente motivación de servicio público, inutilidad y duplicación de trámites, desinformación y pérdida de expedientes, poca racionalidad, autoritarismo, falta de sensibilidad ante los temas empresariales, corrupción, etcétera.

A principios de los setenta dicha percepción negativa se incrementó, influenciada por las ideas de diversos economistas y revistas del campo gerencial, que responsabilizaron a la burocracia de una presupuestación tradicional y una administración financieras ineficientes para garantizar el logro de las metas macroeconómicas y detener el crecimiento del sector público (Barzelay, 2001; Crozier, 1974).

En los ochenta, caracterizados por profundas crisis (fiscal, de endeudamiento, de bienestar, etcétera) y cambios acelerados (de tecnología, condiciones de trabajo, etcétera), la administración pública tuvo que comenzar a cambiar su imagen y hasta su nombre. Así, el término de gestión pública (retomado de la tradición franco-canadiense *gestion publique*)<sup>2</sup> ha venido sustituyendo al desgastado vocablo de “administración pública”, que carga con contenidos poco deseables para el discurso de la modernización o reforma del Estado. A menudo este concepto aparece confundido con el de gerencia pública (traducción habitual de *public management*), que se refiere sólo a la función directiva de este proceso,<sup>3</sup>

<sup>2</sup> Consultar al respecto publicaciones periódicas de la OCDE como “Optique. Le bulletin de la gestion publique”, en [www.oecd.org/puma/index-fr.htm](http://www.oecd.org/puma/index-fr.htm). También documentos del gobierno de Québec, Canadá, publicados en “Modernisation de la gestion publique”, en [www.tresor.gouv.qc.ca](http://www.tresor.gouv.qc.ca)

<sup>3</sup> Paramés (1978: 79) la define como: una función de alto nivel de todas las organizaciones, un proceso de utilización sabia de personas y cosas, y un grupo de personas responsables del gobierno de grupos sociales de muy diversa condición. Guerrero (1998: 20) define a la gerencia pública como el nexo entre las *polícies* y la administración.

con énfasis en la persecución de la eficiencia y preocupada por asuntos ligados a liderazgo y capacidad de mando.

La administración tradicional descrita no dio lugar a desarrollos importantes en materia de evaluación, puesto que su interés estuvo centrado en el control concomitante, directo y detallado de las actividades, en el marco del derecho administrativo y los reglamentos específicos. Sin embargo, en paralelo al proceso reseñado, se había ido gestando un enfoque innovador, que revisaremos a continuación, en el que la evaluación ocupa un papel importante.

### El estudio de las políticas públicas

En la segunda mitad del siglo XX comenzó a destacar el estudio de las políticas (*policies*) públicas. Estas persiguen la resolución de problemas públicos, tomando en cuenta aspectos políticos y también de orden técnico. Se trata de un proceso en el que se entremezclan "...las iniciativas de las organizaciones sociales y las de las instancias gubernamentales" (Aguilar, 2000a: 8) y que intenta responder a una exigencia de racionalidad, legalidad y eficiencia de los gobiernos, sin necesidad de "...comprometerse unilateralmente con alguna específica preferencia conceptual y valorativa sobre el Estado y la acción gubernamental" (p. 9). Para Quade (1975) "...no es sino un complemento, muy poderoso, del buen juicio, la intuición y la experiencia de los «decision makers»" (p. 71).

Una extensa descripción de lo que el concepto puede implicar se encuentra en Subirats (1989, citado en Aguilar, 2000c: 22-23): un campo de actividad gubernamental, un propósito general a realizar, una situación deseada, una propuesta de acción específica, las normas relativas a cada problemática, el conjunto de objetivos y programas de acción del gobierno en un campo de cuestiones; o bien, el producto y resultado de una específica actividad gubernamental, el comportamiento gubernamental de hecho, el impacto real de la actividad gubernamental, el modelo teórico o las técnicas aplicables que sustentan una iniciativa gubernamental, etcétera. La lista demuestra la enorme ambigüedad del concepto y la ausencia de un consenso entre la comunidad de estudiosos que garantice su rigor científico.

En el presente trabajo se define a las políticas públicas como un fenómeno social, administrativo y político específico, resultado de un proceso de sucesivas tomas de posición, que se concretan en un conjunto de decisiones, acciones u omisiones, asumidas fundamentalmente por los gobiernos, mismas que traducen, en un lugar y periodo determinado,

la respuesta preponderante del mismo frente a los problemas públicos vividos por la sociedad civil.<sup>4</sup>

Es así que los gobiernos continúan siendo considerados como los principales actores, únicos que disponen de recursos suficientes para ofrecer bienes y servicios que satisfagan las amplias necesidades sociales. Esto no significa que se niegue importancia a la participación de las organizaciones civiles y sociales o los ciudadanos en general, sino que ésta, si bien se estima deseable, no se considera un elemento imprescindible para definir una política como pública.

En relación a la definición propuesta es necesario aclarar que se habla de un proceso de sucesivas tomas de posición, donde se concatenan planes, programas, proyectos, acciones u omisiones, por lo que se descarta que una decisión aislada, específica, constituya una política; que se hace referencia a problemas públicos como aquellos que, por su relevancia para importantes sectores de la sociedad, no son factibles de resolverse en el ámbito privado; y se considera que los gobiernos no actúan de manera absolutamente coordinada, racional y consistente, por lo que es posible encontrar posiciones diferentes entre las organizaciones que lo componen, y en consecuencia, sólo se pueden identificar tendencias predominantes.

Corresponde señalar que este enfoque de análisis también se ha aplicado a un nivel más desagregado de estudios, el que corresponde a los programas públicos, que permiten una mayor concreción de las políticas en relación a sectores y/o temas transversales específicos. Si bien en el pensamiento anglosajón los términos de políticas y programas pueden aparecer casi como sinónimos (Muller, 1994), en el presente trabajo, se entiende que un programa define con precisión sus objetivos operativos, los medios para alcanzarlos, así como las acciones a realizar en un cierto tiempo; mientras que la política constituye “un conjunto complejo de programas, procedimientos y regulaciones que concurren a un mismo objetivo general” (Kessler *et al.*, 1998: 1, traducción libre), un sistema concatenado y sinérgico de los mismos, que requiere del desarrollo de importantes procesos de coordinación y articulación para lograr el cumplimiento de su misión.

El proceso de políticas puede desagregarse en fases o etapas, enfoque que, aunque no goza de pleno consenso (Dunn, 1981 o Lindblom, 1980; citados en Aguilar, 2000c), resulta dominante entre los estudios de políticas públicas. Esta desagregación ha sido útil para el análisis del proceso, siempre que ha quedado claro que no debe entenderse como una secuencia temporal estricta, sino como una aproximación meramente analítica.

<sup>4</sup>La síntesis presentada, incluida en Cardozo (2001) retoma ampliamente elementos desarrollados en los trabajos de O'Donnell y Oszlak (1976) y Oszlak (1980).



El número de fases del denominado ciclo de las políticas varía con los autores (Laswell, 1971; May y Wildawsky, 1977, citado en Aguilar, 2000b; Brewer y DeLeon, 1983; Anderson, 2000, etcétera), pero todos coinciden en incluir las etapas de aplicación o implementación, evaluación y terminación. El esquema más comúnmente empleado en nuestro medio se conforma de:

- La formulación, que incluye un diagnóstico y definición del problema por parte del organismo estatal, mismo que condicionará el conjunto de alternativas estratégicas de solución que comenzará a generar y evaluar *a priori* para seleccionar la que considere más apropiada. Cabe aclarar que el diseño de políticas, como ya se mencionó, suele estructurarse bajo la forma de programas más concretos, y aún de proyectos específicos, para operacionalizar su ejecución en la siguiente etapa.
- La implementación, que traduce en decisiones operativas y acciones concretas la política antes decidida, de forma de impactar el problema social que le dio origen, procediendo al control o seguimiento de lo formulado.
- Transcurrido el tiempo necesario para poder medir los resultados e impactos alcanzados, podrá procederse a realizar una evaluación *a posteriori* de la política para describir el proceso que se llevó a cabo, explicar sus determinantes y condicionantes y calificar su grado de efectividad. La evaluación permitirá decidir si la política continúa aplicándose y si lo hará con o sin modificaciones, pudiendo dar lugar a la repetición de todas o algunas de las fases señaladas.

En diferentes contextos y periodos históricos, el peso relativo del estudio de políticas se inclinó hacia alguna de las etapas señaladas. En el caso de los Estados Unidos de América (EUA), en los sesenta, el énfasis estuvo en su iniciación; posteriormente, la evaluación resaltó las ineficiencias de las políticas en curso y se volvió predominante. Su análisis identificó serios problemas de implementación, con lo que esta etapa cobró importancia en la segunda mitad de los años setenta, acompañada en los casos de mayor fracaso, por la consideración de la etapa de terminación. En los ochenta, se produjo el redimensionamiento del Estado y se regresó al interés por el diseño de políticas, capaces de aumentar la eficiencia de la gestión pública (Aguilar, 2000b). Hoy asistimos a una renovación del interés por la etapa de evaluación en el contexto de las reformas que revisaremos en el punto siguiente.

Por otra parte, las diferentes ciencias que participan del enfoque inter y multidisciplinario de las políticas públicas, también han concentrado su

interés en el análisis de algunas fases en detrimento de las restantes: “Es así que los politólogos han preferido estudiar la formación de la agenda de gobierno, los administradores públicos la implementación y los economistas se han dedicado enteramente a la construcción y selección de las opciones de política” (Aguilar, 2000c: 19). Podríamos agregar también que los sociólogos se interesan por la gestación de los problemas en la sociedad, aunque en los últimos años, en Francia, han sustituido a los economistas en la realización de evaluaciones. Estas especializaciones han logrado importantes aportaciones al conocimiento, pero también han tenido resultados no deseados como un economicismo que poco recupera los análisis de impacto social y político en los procesos de formulación de políticas o un fuerte abandono de la evaluación, objeto principal de preocupación en el presente trabajo.

Laswell (1951), iniciador de este campo de estudios, definió a las ciencias de las políticas como las “...disciplinas que se ocupan de explicar los procesos de elaboración y ejecución de las políticas...” (citado en Aguilar, 2000a :47), con base científica, interdisciplinaria y al servicio de gobiernos democráticos. En un trabajo posterior (1971) agregaría que era posible diferenciar “el conocimiento «del» proceso de decisión y «en» el proceso de decisión” de políticas, refiriéndose a las habilidades profesionales necesarias para participar en la toma de decisiones públicas y las habilidades científicas requeridas para contribuir a la invención de una teoría y práctica confiables (p. 117).

La diferenciación anterior daría lugar a dos corrientes: por un lado, quienes se autoidentificaron fundamentalmente bajo el rótulo de *Policy Analysis*, interesados exclusivamente en el momento normativo-técnico de análisis y diseño de la política, desde un enfoque unidisciplinario; y, por otro, los partidarios de las *Policy Sciences*, preocupados sólo por entender los procesos decisorios, como suceden en los hechos, desde una perspectiva multi o interdisciplinaria.

En este trabajo se ubica a la evaluación de políticas y programas públicos como parte del conocimiento “del” proceso, en la medida en que busca analizar científicamente los resultados de los procesos de decisión e implementación. La misma puede ser desarrollada por el propio gobierno o el Congreso, pero también por organizaciones civiles que representen los intereses de la sociedad.

El tratamiento de las políticas públicas aparece también surcado por diferentes enfoques que podríamos esquematizar de la siguiente manera:

Un enfoque de racionalidad total, que supone un decisor individual o diversos decisores organizacionales que actúan como bloque unificado,

disponen de información completa y una capacidad ilimitada de cálculo que les permite revisar exhaustivamente todas las alternativas disponibles y encontrar siempre la mejor solución a los problemas públicos (optimización). Este enfoque se caracteriza por ser normativo y fuertemente cuantitativo.<sup>5</sup>

Un segundo acercamiento desde el análisis del proceso organizacional que pone el énfasis en la explicación de la decisión tomada con base en las rutinas y procedimientos estándares vigentes en las organizaciones, parte de la idea de una racionalidad limitada que lleva a alcanzar decisiones satisfactorias, en lugar de óptimas.<sup>6</sup>

Existe otro enfoque fundamentado en el proceso político incrementalista en que cada actor se mueve por sus propios intereses y negocia con los demás sin compartir sus fines, por lo que se considera que la optimización no es posible (no se cuenta con información completa ni existe capacidad ilimitada de cálculo) ni deseable (no existen objetivos compartidos).<sup>7</sup>

Un cuarto enfoque aporta el conjunto de elementos no racionales del proceso decisorio, en que se recuperan elementos como intuición y experiencia desde perspectivas psicosociológicas, antropológicas y filosóficas.<sup>8</sup>

Finalmente podemos encontrar mezclas de los distintos enfoques, como el planteado por A. Etzioni, quien critica al modelo racional por las razones ya planteadas, pero también al incremental por su tendencia conservadora, que descarta los cambios radicales. Propone una recuperación selectiva de elementos de los distintos enfoques anteriores.

En el presente trabajo de evaluación *a posteriori* optamos por esta última posición al considerar que las políticas y programas públicos atraviesan y provocan diversos procesos, con distintos tipos de racionalidad no exhaustiva:

1. políticos, en la definición de los valores y objetivos a perseguir para alcanzar algún nivel de satisfacción de la ciudadanía que permita el mantenimiento o incremento de la legitimidad de los actores sociales (incluido el gobierno);
2. técnicos, de ajuste de medios a fines en la identificación y selección de políticas y programas a aplicar para el logro de los anteriores (racionalidad administrativa instrumental, de tipo limitada); y

<sup>5</sup> Sus orígenes los podemos encontrar desde la propuesta cartesiana, desarrollada en *Teoría de Juegos* por Von Neumann y Morgenstein y en *Teoría de la Estrategia* por Von Clausewitz.

<sup>6</sup> El enfoque es iniciado con los trabajos conductistas de Herbert Simon. *El comportamiento administrativo* (1947) y *Models of Man, social and rational* (1957), y continuado por Cyert, March, Olsen, Cohen, etcétera.

<sup>7</sup> Charles Lindblom retoma la idea de incrementalismo planteada por el enfoque anterior, atribuyéndole una valoración diferente en su famoso artículo "The science of muddling trough", aparecido en 1959.

<sup>8</sup> Citado a pie de página por Allison (1971) y desarrollado en Steinbruner (1974).

3. sociales, vinculados a la posible participación de la ciudadanía y las organizaciones de la sociedad civil en los procesos previos y al imprescindible análisis de los impactos provocados en la población por las estrategias aplicadas.

También se ha cuestionado si el estudio de las políticas públicas constituye una disciplina o un campo de acción multidisciplinario (muy vinculada con la misma discusión que enfrenta la administración pública como disciplina), científico o no científico. Dada la trascendencia de esta discusión, en la que no es posible profundizar en este trabajo, se le considerará un campo profesional en el que concurren múltiples disciplinas, como sucede con la medicina. Se pueden hacer evaluaciones que apliquen el método científico para describir y explicar las políticas implementadas y sus efectos. También es posible que la formulación se base en diagnósticos científicos de la situación, pero el análisis de toma de decisiones está fundado en la previsión de posibles consecuencias futuras, marcadas por la incertidumbre, donde las categorías de “verdadero” y “falso” pierden sentido y se sustituyen por otras como “probable” o “factible”, en cuyo trabajo de prospectiva influyen también aspectos como experiencia, sensibilidad, intuición, etcétera, además del conocimiento científico, lo que dificulta la definición de su estatus.

Actualmente, Majone (1996) identifica tres rasgos relativamente nuevos que caracterizan las preocupaciones de las políticas públicas: el redescubrimiento de la necesidad de eficiencia, el reconocimiento de la importancia estratégica de la credibilidad política, y el aumento en la delegación de poderes a unidades independientes para la hechura de políticas; todos ellos estrechamente vinculados al tema de la evaluación.

### El New Public Management

En las dos últimas décadas, y con la intención de resolver múltiples problemas burocráticos, dio inicio un esfuerzo de reforma del Estado y modernización de la administración pública, impulsado por el enfoque denominado New Public Management (NPM), que ha dado un nuevo y fuerte impulso a los procesos de evaluación.

Este enfoque, traducido con dificultades en nuestro medio como Nueva Gestión o Nueva Gerencia Pública, se apoya en el Neoinstitucionalismo Económico y el Public Choice, así como en experiencias prácticas calificadas como exitosas en varios países sajones, particularmente en Nueva Zelanda, Reino Unido y Australia, que introdujeron reformas presupuestales dirigi-

das al control del gasto y el logro de resultados, acompañadas de importantes modificaciones legales y cambios institucionales, así como de un fuerte proceso de privatización y de rendición de cuentas.<sup>9</sup> Se trata de un movimiento que ha recibido un decidido apoyo de la Organización para la Cooperación y el Desarrollo Económicos (OCDE), el Fondo Monetario Internacional (FMI) y el Banco Mundial (BM).

Resulta importante subrayar la idea de que gestión y gerencia pública son campos preexistentes al nacimiento del NPM y que, a su interior, existen enfoques distintos y hasta opuestos a éste en algunos aspectos, como el defendido por Bernardo Kliksberg (1989 y 1994) que no admite la crítica generalizada a la ineficiencia del sector público y lejos de tender a equiparar la administración pública con la privada, enfatiza sus diferencias, incluso las de la gerencia social frente a la gerencia económica, con su particular papel de lo político, la preocupación por la articulación (no subordinación) de las políticas económicas y sociales, los procesos de negociación, la priorización de los intereses de las grandes mayorías empobrecidas, la aplicación del enfoque de políticas públicas, el mayor énfasis en la participación como mecanismos de control social y el desarrollo de programas sociales masivos.

El NPM no constituye un nuevo paradigma, ya que al respecto no existe consenso entre los científicos del campo (Kuhn, 1971), ni un programa sistemático de investigación sobre el mismo (Lakatos, 1989), ya que no contiene un conocimiento ordenado y sistematizado en la materia, sino una combinación de estudios relacionados, con una orientación común: el incremento de la eficacia, economía, eficiencia (lema de la triple E utilizado por Margaret Thatcher) y calidad del servicio. El mismo ha sido definido como un campo de discusión sobre las intervenciones del gobierno en materia de políticas, tanto en términos de análisis sistemático como de gerencia práctica, enfocada a los procesos políticos y organizacionales así como también a sus contenidos sustantivos (Barzelay, 2001); como un movimiento genérico que cruza diferentes sectores de las políticas y fronteras nacionales (Hood y Jakson, 1997); o se ha reconocido a su interior un conjunto de modelos diferenciados como el manejo de la eficiencia, la reducción y descentralización del Estado, la búsqueda de la excelencia, o la orientación al servicio público (Ferlie *et al.*, 1996).

<sup>9</sup>Guerrero (1999: 129) plantea que Michel Messenet, director general de la Función Pública en Francia, empleaba ya, en 1975, terminología e ideas del NPM como orientación al mercado, privatización, descentralización, preocupación por los costos y la eficiencia administrativa, consideración del ciudadano como cliente, etcétera, en su libro *La Nouvelle Gestion Publique: pour un État sans bureaucratie*, París, Editions Hommes et Techniques. A su vez Messenet reconoce haberse inspirado en los trabajos de un administrativista privado, Octave Gelinier, quien hizo el prólogo de su libro y sería, a su vez, autor de *Le secret des structures compétitives*, publicado en 1967. En el contexto del Reino Unido, Guerrero señala que fue recién en 1993 que apareció una obra similar. *Public Sector Management*, de N. Flynn.

Entre las diversas estrategias que el NPM propone para lograr los objetivos de su agenda internacional encontramos: reducción del tamaño de la burocracia y selección de lo que el gobierno debe hacer, descentralización, separación entre la formulación y la implementación de la política, jerarquías aplanadas, facultación y responsabilización, coordinación de las interdependencias, competencia entre organizaciones públicas, generación de valor para el cliente, cooperación público-privado, enfoque de gobernanza participativa (de consumidores, empresarios y empleados públicos), distinción entre el financiamiento y la prestación de servicios públicos, apertura a la competencia, privatización, utilización de técnicas de gerencia privada, flexibilización, simplificación de trámites, desregulación, democratización, enfoque proactivo de problemas, énfasis en costos, calidad, libertad de elección,<sup>10</sup> satisfacción y empoderamiento del cliente, cultura orientada a resultados, focalización de programas, evaluación del desempeño, utilización de incentivos de mercado, transparencia contable, rendición de cuentas, etcétera.

Para Abate (2000) su aplicación requiere de una evolución general de la sociedad que garantice: 1. una amplia base profesional suficiente para gerenciar en forma autónoma la gestión pública (competencias y capacidades de iniciativa de los agentes); 2. disposición de redes informáticas y de tratamiento de datos; y 3. capacidad de los sistemas de información de medir los costos y los productos de las acciones de la gestión para evaluar su calidad y su impacto en la sociedad.

En el contexto latinoamericano, el consejo científico del CLAD, aprobó en octubre de 1998 el documento “Una nueva gestión pública para América Latina”, que explicita su posición oficial sobre la reforma gerencial. Asume que este tipo de reformas ofrece las mejores posibilidades en términos de democracia y progreso para la región, pero también plantea el desafío de compatibilizar las tendencias mundiales con sus especificidades (necesidad de consolidar la democracia, de recuperar el crecimiento económico y de reducir la desigualdad social). Este documento propone: profesionalización de la alta burocracia, transparencia de la administración pública, descentralización de la ejecución de servicios, desconcentración de las actividades del gobierno central, orientación a resultados, mayor autonomía compensada con nuevas formas de control, distinción entre actividades exclusivas del Estado y otras del espacio público no estatal, orientación de los servicios hacia el ciudadano-usuario, democratización del poder público y mayor responsabilización de políticas y funcionarios.

<sup>10</sup> Basada en la noción de “soberanía del consumidor” de Ludwig Von Mises.

Para su puesta en práctica el CLAD señalaba como prerequisite completar la reforma burocrática weberiana y redefinir las relaciones entre el Estado y la sociedad, construyendo una esfera pública que involucre a esta última en la gestión de las políticas públicas.

Uvalle (1998, 2001, 2002 y 2003) coincide con el CLAD en que las reformas deben tener dos objetivos: gestión democrática y gobierno eficiente. El último, que caracteriza como un “gobierno facilitador”, se define por una gestión conjunta con los ciudadanos para la atención de los problemas comunes y por el control democrático del rendimiento institucional. Considera que la recuperación del espacio público por la sociedad reduce las posibilidades de volver a contar con “Estados magnificados”.

En relación con los resultados alcanzados por las reformas, un informe producido por el comité de gerencia pública de la OCDE (1995) señalaba que era muy pronto para juzgar el impacto de los cambios realizados, pero también se reconocía una lamentable escasez de evaluaciones rigurosas de los mismos que permitiría construir mayor confianza y promover una cultura basada en el enfoque gerencial. Se aseguraba que las experiencias “sugerían” éxito en el control del gasto público y que se disponía de “evidencia considerable” en el caso de agencias individuales para “hacer más con menos”, incrementar su eficiencia y la calidad de sus servicios. En cambio, se reconocía la dificultad y complejidad de medir la efectividad de las reformas introducidas, por lo que muchas conclusiones se calificaban de “tentativas” y exigían continuar manteniendo como supuestos del enfoque los referidos a que la utilización de incentivos permitirá mejorar los resultados o que la descentralización mejorará el nivel de logros al reducir los costos asociados con los controles centrales.

Nueva Zelanda, el caso emblemático del NPM, tampoco cuenta con evaluaciones consistentes sobre el éxito de sus reformas: si bien logró éxitos financieros y macroeconómicos de gran magnitud, estos aparecen acompañados de algunos resultados no deseados como un fuerte incremento en el desempleo.<sup>11</sup>

En el caso de América Latina (AL), Cohen (2001) plantea que, a pesar de la importancia de los mecanismos de descentralización, externalización

<sup>11</sup> Aumento del crecimiento del PIB del 77 por ciento, del PIB per cápita de 160 por ciento, de la productividad del trabajo del 400 por ciento, disminución de la inflación del 57 por ciento, del déficit en cuenta corriente del 24 por ciento etcétera, en el periodo de 1983-1997, comparado con el de 1973-1983, aunque estos no han permitido recuperar los niveles de los mismos indicadores que el país gozó en el periodo 1963-1973. Los anteriores se acompañaron de un aumento de la tasa de desempleo del 267 por ciento (pasó de 1.8 a 6.6 por ciento), en la comparación de los mismos periodos y una disminución del 17 por ciento del total de empleos, en el sector público, en 1989-1997.

y focalización implementados en los programas sociales como parte de las reformas en curso, no existe claridad sobre los beneficios de su aplicación. Trabajos como el de Spink (2001) o el de Longo y Echebarría (2001) argumentan a favor, en términos de posibilidades, pero también en contra, recogiendo las limitaciones del enfoque.

Esta falta de pruebas contundentes ha llevado a que los estudiosos del NPM se dividan en un grupo que defiende teóricamente las bondades de la propuesta (Barzelay, 1992 y 2001; Osborne y Gabler, 1994; Uvalle, varios años; Osborne y Plastrik, 1998; Hood, 1994; OCDE, 1995) y otro que la cuestiona por su ideología, inconsistencias o pretensión de aplicación universal (Pollit, 1993; Christensen y Laegreid, 2001; Guerrero, 2004; Brugué, 2004). Thoenig (2002) señala que la marcha triunfal del NPM ha comenzado a frenarse en el contexto del BM, el FMI, la OCDE, Reino Unido, Nueva Zelanda, etcétera.

En el presente trabajo, más que involucrarnos en la polémica anterior, nos interesa resaltar el nuevo impulso que el NPM ha dado a la evaluación como herramienta de buena gestión. La OCDE (1995) señalaba cinco características fundamentales a desarrollar por los gobiernos en la siguiente década; dos de ellas se referían a la evaluación continua de la efectividad de las políticas y reformas y al impulso de un enfoque de gobernanza, más participativo.

Es que más allá de la necesidad que históricamente ha tenido la administración pública centralizada de contar con retroalimentación sobre sus logros para corregir y rediseñar sus acciones, los cambios introducidos en materia de descentralización, la delegación de mayores facultades, flexibilidad y libertad (disminución del control jerárquico) a los gerentes para el logro de los objetivos comprometidos, la participación de agencias no gubernamentales, etcétera, incrementan el requerimiento de una cultura orientada a resultados<sup>12</sup> que necesita ser evaluada, tanto para la aplicación de incentivos positivos y negativos, como para la transparencia contable y la rendición de cuentas, insumos básicos de la gobernanza participativa. Así, la traducción de los objetivos estratégicos del gobierno en metas operativas y la verificación de su cumplimiento constituyen la verdadera esencia de la gerencia pública (OCDE, 1995). También la Unión Europea (UE) o la Organización de las Naciones Unidas (ONU, 1994), menos permeables al enfoque del NPM, han insistido en la necesidad de la evaluación de los programas y el desempeño de los funcionarios.

<sup>12</sup> Esta orientación hacia el logro de resultados fue retomada de antecedentes como el Planning-Programming-Budgeting System (PPBS) de los EUA, la Rationalisation des Choix Budgetaires (RCB) francés o el enfoque de administración por objetivos.



Por las razones anteriores, el NPM refuerza los sistemas de información y evaluación de los resultados de la gestión pública (políticas, programas, organizaciones, funcionarios), focalizándolos hacia la medición de la eficiencia y economía de operación, la calidad de los servicios y las relaciones costo/efectividad. En algunos países como Australia los resultados se miden en conjunto, pero en otros como Nueva Zelanda o el Reino Unido se hace una clara diferenciación entre resultados en tanto productos o servicios brindados (*outputs*) y sus impactos (*outcomes*), siendo los primeros, responsabilidad de los gerentes, y los segundos, de los ministros. Bélgica, Francia o España han enfatizado el análisis de la calidad de los servicios.

Los sistemas de evaluación requieren de dos actividades complementarias: 1. el diseño de indicadores apropiados de seguimiento continuo del desempeño y sus correspondientes estándares (con base en logros previos, logros de otras organizaciones similares, etcétera); y 2. la realización de evaluaciones periódicas en profundidad. La primera ofrece señales sobre la posible existencia de problemas y la segunda se aplica para su verificación, explicación y valoración, o sea, intenta establecer relaciones causa-efecto entre los factores explicativos y el nivel de logro alcanzado por el programa, mismos que permitan tomar decisiones sobre su continuación, conclusión o modificación en aspectos de financiamiento, delegación de facultades y/o rediseño de sus contenidos. De esta forma las actividades de evaluación se integran con las de planeación-presupuestación, dando lugar a una gerencia por resultados.

Dicha evaluación “no es una ciencia exacta”... “este punto no siempre ha sido entendido por los legisladores e incluso por los ministros”... “quienes pueden hacer un uso acrítico de la información sobre resultados en la creencia de que todo logro puede ser plena y precisamente cuantificado” (OCDE, 1995: 104, traducción libre). Dicha información debe ser considerada en el marco de su definición conceptual, metodológica y empírica y puede haber enfrentado dificultades de distintos tipos: la definición poco operativa de los objetivos planeados, la multiplicidad y conflicto entre objetivos, la complejidad de la información analizada, la difícil verificación de relaciones de causalidad, la incompatibilidad de tiempos para su realización y su utilización en el ciclo presupuestal, la insuficiencia de recursos debida a las reducciones del gasto público, la falta de independencia de los evaluadores, etcétera.

En el mismo sentido, Abate (2000) señala la dificultad de gerenciar los servicios no comerciales del Estado, en pos de la eficacia en la producción de servicios colectivos para beneficio de la sociedad, en comparación con la simple búsqueda de rentabilidad en una empresa privada. En cuanto a la evaluación de la gestión por resultados la considera una disciplina a crear, pues

se requiere de nuevas formas de pilotaje capaces de manejar la multiplicidad de factores que determinan un cierto nivel de logro, muchos de ellos relativos a la sociedad y no sólo provocados por las acciones de la gestión pública. Duran y Thoenig (1996) agregan otro nuevo factor a esta complejidad: las características multiorganizacionales de los programas públicos, que obligan a la evaluación a superar los límites de las organizaciones individuales como unidad de análisis.

A pesar de las dificultades anteriores, la información sobre logros debe ser usada en los procesos decisorios; para ello debe ser reportada a las personas que la requieren y en forma diferente según sus responsabilidades (ministros, jefes ejecutivos, legisladores, etcétera). "...la obtención de ganancias de logro reposa sobre la comprensión por los actores de los procesos en curso y sobre las laxitudes que les son dadas para actuar" (Abate, 2000: 53, traducción libre). Además, el autor señala que la orientación a resultados no se logrará plenamente si la evaluación se limita a alimentar el interior de la gestión pública; considera fundamental el diálogo con los ciudadanos y el parlamento como un medio poderoso para estimular su logro.

Por otra parte, la calidad de la evaluación de resultados puede ser mejorada mediante la realización de metaevaluaciones, o sea, "evaluaciones de las evaluaciones".

### Aportes sociológicos

En los estudios clásicos de ciencias políticas, el Estado y el gobierno se definen por la dominación y el poder que ejercen sobre la sociedad y la economía en un cierto territorio; en cambio, en la sociología de la acción pública se definen como un proceso de coordinación de actores, grupos sociales e instituciones para alcanzar los fines propios, discutidos y definidos colectivamente en contextos fragmentados e inciertos (Mayntz, citada en Commaillies y Jobert, 1998). Este proceso resulta más cercano al concepto de *governance* (traducido en español como gobernanza o gobernancia), o capacidad para gobernar, vista a través de la producción de redes de actores gubernamentales y cívicos en el desarrollo de políticas públicas. En su estudio se retoman las ideas planteadas por el Neoinstitucionalismo Sociológico y la Teoría de la Contingencia, en torno a la importancia de la cultura y el marco institucional de las organizaciones para explicar las dificultades que enfrentan los procesos de cambio, especialmente los inducidos.

La sociología política maneja el concepto de regulación social para referirse al modo de coordinación de las diferentes actividades o relaciones

entre actores, la asignación de recursos relacionada con estas actividades y actores y la estructuración de los conflictos, su prevención y resolución; así como los procesos de creación y aplicación de reglas para asegurar la coexistencia, en un mismo territorio, de individuos con intereses divergentes, incluso contradictorios y potencialmente conflictivos. Se vincula al tema de la integración social y trata de la acción colectiva organizada, bajo formas multipolares, poco o no formalizadas, y no necesariamente coercitivas. La diversidad de la última provoca que su coherencia sea problemática, sus formas variadas y cambiantes y su efectividad incierta; lo que obliga a pensar en formas renovadas de intervención social.

Para esta disciplina, la regulación social se efectúa mediante tres mecanismos complementarios:

- regulación por el Estado, en que éste estructura los conflictos, distribuye recursos y coordina actividades y grupos, para lo que requiere capacidad de control y sanción;
- regulación por el mercado, que permite la organización de intercambios en el capitalismo; y
- regulación por la cooperación o reciprocidad, basado en valores, normas, identidades, intercambios movidos por la confianza, solidaridad entre miembros de la familia, comunidad, etcétera (Commailles y Jobert, 1998).

La sociología de las organizaciones, por su parte, ha superado el énfasis casi exclusivo en la organización formal, para recuperar el análisis de la informalidad, ya planteado por Simon (1972). La organización se ha vuelto un sistema social en funcionamiento concreto, producto de un trabajo de construcción de todos sus miembros, en que lo formal y lo informal se han podido conciliar mediante el estudio del concepto de poder. Se trata de analizar a la organización para identificar "...sus actores, intereses, recursos, restricciones, comportamientos, en el contexto específico de los juegos relacionales que estructuran la organización" (1998: 39-40, traducción libre), relativizando sus posiciones y poderes específicos, al mismo tiempo que descubriendo las reglas implícitas (de segundo orden o no escritas) que compatibilizan intereses heterogéneos y son respetadas por los actores. La racionalidad de las decisiones individuales se descubre después de describir y explicar el juego colectivo.

Este enfoque de las organizaciones se ha extendido a la comprensión de los sistemas de acción, o sea, las relaciones de interdependencia e intercambio

entre individuos y grupos que intervienen alrededor de un problema, lo que nos permite aplicarlo al análisis de las políticas y programas públicos, sustituyendo el enfoque *top down* de la sociología del Estado todopoderoso por un acercamiento *bottom up*.

También aporta el concepto y funcionamiento de las estructuras cognitivas de Bourdieu, que permiten describir representaciones, percepciones y visiones del mundo que tienen los individuos, así como las estructuras que determinan su formación y transformación (Chauviré y Fontaine, 2003). Las estructuras cognitivas, a veces denominadas estructuras mentales, hacen que los individuos acepten el mundo como algo natural, provocando una especie de “sumisión” cognitiva en lugar de una verdadera actividad cognitiva. Constituyen la prolongación, a nivel individual, de las estructuras objetivas de los diferentes “campos” sociales.

La aplicación de las aportaciones de la sociología de las organizaciones ha sido muy productiva en el contexto francés, en que el Estado ha dejado de disponer de importantes recursos financieros y jurídicos, ya no controla los programas que, si bien son decididos por el poder central, son aplicados por las administraciones locales. Ahora ya no se trata de producir infraestructura o de brindar servicios públicos, sino de resolver problemas difusos, poco estructurados, a los que no se adaptan los mecanismos burocráticos. En este contexto, la coordinación y la negociación se consideran fundamentales para la articulación de fines colectivos a partir de intereses de múltiples actores diferenciados y se impone la cooperación y comunicación en el marco de redes de solución de problemas.

Un ejemplo importante de esta aportación sociológica lo constituye el “modelo de las regulaciones cruzadas”, desarrollado para comprender el cambio de los asuntos públicos franceses en los sesenta-ochenta, centrándose en el análisis empírico de las estrategias de los actores, interpretando los juegos colectivos a los que responden las anteriores y los modos de regulación que gobiernan estos juegos y constituyen la características más pertinentes de los sistemas, que conforman redes más o menos organizadas de relaciones humanas (Crozier y Thoenig, 1975).

Más recientemente, la sociología de las organizaciones ha orientado el estudio de la gestión pública territorial a través del modelo de institucionalización de la acción colectiva (Duran y Thoenig, 1996 y Duran, 1999), en el marco de las transformaciones sufridas por el Estado francés (debilitamiento de su poder, descentralización, aparición de actores locales y grupos de interés, multiplicación de las externalidades, etcétera). Su aplicación a nivel territorial permite descubrir esquemas de cooperación informal que explican la solidaridad entre

actores que ocupan puestos, niveles jerárquicos e inserciones institucionales diferentes (Thoenig, 1994).

En materia de evaluación este enfoque sugiere analizar: 1. la forma en que diversos actores con intereses divergentes logran volverlos compatibles en la organización, a pesar de todas sus contradicciones y conflictos; 2. las normas implícitas que regulan sus relaciones recíprocas; 3. las consecuencias que estos arreglos inducen; y 4. el grado en que dichas consecuencias son “pagadas” por grupos internos a la organización social o externalizadas sobre terceros. Esta última manifestación permite superar el corte de los estudios exclusivamente volcados hacia el interior de las organizaciones, al analizar la forma en que los actores externos afectan o son afectados por éstas. También recomienda tomar en cuenta las diferencias culturales<sup>13</sup> en los procesos de adaptación de tecnología administrativa desarrollada fuera del país (en materia de evaluación, por ejemplo, los estudios experimentales producidos en los EUA).

### Enfoques emergentes

Además de las críticas ya mencionadas que se han venido formulando al NPM durante las dos décadas de su aplicación, en la actualidad parece percibirse una mayor tendencia a proponer modelos alternativos o, al menos, ajustes importantes en su aplicación.

Brugué (junio de 2004), después de criticar como debilidad principal del NPM su sesgo hacia la competencia y el mercado, mismo que considera que lejos de transformar a la administración pública tradicional, ha reproducido sus rasgos eficientistas, propone una alternativa que denomina “Administración Pública Deliberativa”, que se percibe muy cercana al gobierno por discusión sugerido por Guerrero (1999).

Dicho modelo, sin abandonar la eficiencia como objetivo instrumental, se enfocaría al objetivo sustantivo de la modernización administrativa en una forma cercana a la propuesta sociológica: resolver los conflictos y responder a las demandas del entorno social. Esto exige desarrollar una administración basada en la confianza y el diálogo, para lo cual se requieren nuevas formas y estilos de liderazgo.

<sup>13</sup> Una caracterización de la diversidad de las culturas nacionales ha sido realizada por Geert Hofstede, en *Cultures and organizations. Software on the mind*, McGraw Hill Internacional, 1991, con base en cuatro dimensiones: desigualdad social o índice de distancia al poder, grado de individualismo-colectivismo, rechazo-aceptación de la incertidumbre, y diferenciación y valores asociados al género.

En el caso de AL y en los dos últimos años, se percibe un viraje importante en la posición del CLAD respecto de las reformas inspiradas en el NPM. Su consejo directivo convocó, en 2003, a la V Conferencia Iberoamericana de Ministros de Administración Pública y Reforma del Estado para debatir el tema de los consensos políticos y sociales necesarios, por considerar que muchas decisiones de reforma “han sido impuestas por urgencias” que han impedido su construcción, lo que dio lugar al consenso de Santa Cruz.

En 2004, el documento “Democracia en América Latina: hacia una democracia de ciudadanos y ciudadanas” del Proyecto de Naciones Unidas para el Desarrollo (PNUD) confirmó algunos datos anticipados en 2003 por Latinbarómetro en el sentido del limitado avance de la democracia en la región y el debilitamiento de la gobernabilidad.

En el mismo tenor, el documento introductorio a la VI Conferencia Iberoamericana de Ministros de Administración Pública y Reforma del Estado convocada por el CLAD, señala que: “...gran parte de las condiciones negativas para la gobernabilidad democrática provienen de que las reformas estructurales que hasta ahora se diseñaron y ejecutaron... han sido orientadas por las necesidades de un determinado enfoque económico, cuyo resultado ha sido la disminución del apoyo a la democracia” (Fernández Toro, junio de 2004: 1).

Frente a esta situación, el CLAD propone redefinir las jerarquías entre economía y política para dar prioridad a la democracia frente al mercado: vigencia del estado de derecho, efectividad de los derechos políticos, civiles y sociales y participación efectiva de la ciudadanía en las decisiones sobre las cuestiones sustantivas de la sociedad. Esto se refuerza en Bresser Pereira *et al* (2004), quienes enfatizan la necesidad de que la ciudadanía participe como un actor político clave y se produzca un mayor control social sobre los procesos de reforma.

En lo que respecta a la evaluación, Zaltsman (junio de 2004) concluye, con base en el análisis del trabajo desarrollado por Cunill y Ospina (2003), que en AL hay casos nacionales en que la evaluación de resultados de la gestión resultó muy afectada por los valores y principios del NPM, pero también otros en que la participación de la ciudadanía y la profundización de la democracia representativa jugó un rol igualmente importante. Además, no identifica indicios de incompatibilidad o conflicto que impidan a la evaluación diseñada la consideración de valores diferentes como la democracia y la equidad social.

## Comentario final

Después de 50 años de relativo desinterés de la burocracia tradicional por la evaluación, ésta fue ganando terreno con el desarrollo de los estudios de políticas públicas e incrementando su importancia en el marco de las estrategias promovidas por el NPM. Los enfoques sociológicos y emergentes tampoco han cuestionado a la evaluación misma, sólo han tratado de infundirle valores y objetivos diferentes a los impulsados por el NPM.

El siguiente capítulo profundizará en el abanico de aproximaciones teóricas y conceptuales presentes en el desarrollo actual de la evaluación de políticas y programas públicos.

## Capítulo 2

# La evaluación desde el punto de vista teórico

### Introducción

DESPUÉS de esbozar la evolución del papel jugado por la evaluación en la administración pública, el presente capítulo tiene por objetivo sintetizar las diversas contribuciones teóricas que a la fecha se han realizado en materia de evaluación de políticas y programas públicos, especialmente en su enfoque *a posteriori*, ofreciendo un panorama general de los conceptos, objetivos, enfoques, actitudes y destinatarios de la misma, y resaltando la importancia del contexto político-organizativo en que se desarrolla.

### Concepto de evaluación

Como tantos otros conceptos de las ciencias sociales, podemos encontrar casi tantas definiciones de evaluación como autores han escrito sobre ella. Sus diferencias suelen girar alrededor de énfasis distintos atribuidos a los elementos que componen el concepto en estudio, por lo que podemos hablar fundamentalmente de un conjunto de definiciones que se complementan, más que de concepciones opuestas. M. Weber habría sido uno de los primeros en definir a la evaluación como "...la apreciación (*bewertung*) práctica de un fenómeno sobre el cual nuestra actividad puede ejercer una influencia al adoptar respecto del mismo una actitud de aprobación o desaprobación" (citado en Duran y Monnier, 1992, traducción libre).

Sintetizando las aportaciones de múltiples autores (Guba y Lincoln, 1981; Nioche, 1982; Stufflebeam y Shinkfield, 1987; Guba, 1987; Rossi y Freeman, 1993; Mény y Thoenig, 1995; Dye, 1995, citado en IDR, 2000; Weiss, 1998; Duran, 1999; Banco Mundial, 2004a), es posible considerar a la evaluación como una investigación aplicada, de tendencia interdisciplinaria, realizada mediante la aplicación de un método sistemático, cuyo objetivo es conocer, explicar y valorar una realidad, así como aportar elementos al proceso de toma de decisiones, que permitan mejorar los efectos de la actividad



evaluada. Se resalta la importancia atribuida a la emisión de un juicio valorativo sobre dicha actividad, lo que también es compartido por Scriven (1967) y Aguilar y Ander-Egg (1992).

Algunos autores consideran (Monnier, 1995; Solarte, 2002; Díaz y Rosales, 2003) que definiciones como la presentada enfatizan su carácter instrumental, por lo que ensayan definiciones de corte más sociológico y político, referidas a la evaluación como un proceso social e histórico de construcción de consensos y acuerdos, transformación de las ideas y los conocimientos, basado en el diálogo, con enfoque metodológicos múltiples y distintos del discurso racional de la teoría de las decisiones, que contribuye a modificar los programas públicos en forma lenta y marginal. Resaltan la importancia de que la evaluación coadyuve al logro de una democracia efectiva, a través de la rendición de cuentas del gobierno a los ciudadanos y de su posible participación en el proceso.

Resulta importante diferenciar a la evaluación del mero control o fiscalización de la gestión, que se enfoca a la utilización de los recursos presupuestales, y de la auditoría tradicional, centrada en verificar la adecuación de los procedimientos legales y administrativos.

La realización de una evaluación implica un conjunto de actividades, cuya secuencia varía según el tipo de evaluación a realizar:

- a) identificar los efectos provocados por la acción, programa, política o evaluación a evaluar y los costos incurridos;
- b) aplicar una escala de medición a los efectos identificados (nominal, ordinal, de intervalo o razón);
- c) comparar la medida lograda con otra que sirva de parámetro para la evaluación (ejemplo: la medida de efectos logrados en periodos previos por la misma organización, política o programa; en el periodo actual por organizaciones, políticas o programas similares; la determinada en los planes organizacionales como meta esperada, etcétera), fundamentalmente, en el caso de los programas sociales, para constatar si hubo mejoramiento de las condiciones de bienestar de la población;
- d) explicar los resultados comparativos encontrados, en función del diseño y las condiciones de aplicación del programa;
- e) emitir un juicio de valor que califique, en conjunto, las actividades realizadas, los servicios brindados, sus efectos y su impacto global, y
- f) efectuar las recomendaciones necesarias para enfrentar los problemas detectados y aprovechar las fortalezas del programa, de manera de contribuir a un mayor logro de sus objetivos.

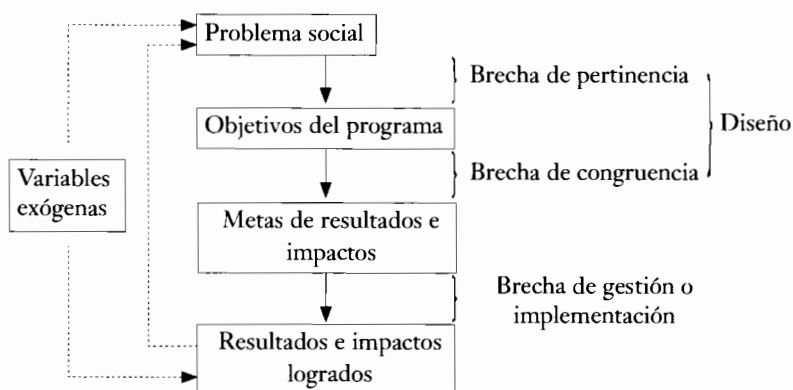
Las acciones evaluadas pueden vincularse a diversos objetos de evaluación: personas, organizaciones, políticas, planes, programas, proyectos o evaluaciones. En el presente trabajo el objeto de la evaluación lo constituirán las evaluaciones aplicadas a las políticas y programas públicos, especialmente los que persiguen objetivos sociales.

Los anteriores pueden ser muy diversos (combate a la pobreza, educación, vivienda, salud, medioambiente, empleo, servicios básicos, etcétera), e implicar conocimientos específicos de campos variados.

### Justificación de la evaluación

Tal cual puede verse en el diagrama 1, cualquier problema social es afectado por múltiples causas contextuales y por el o los programas que se hayan diseñado para tratar de solucionarlo.

DIAGRAMA 1  
JUSTIFICACIÓN DE LA EVALUACIÓN



Fuente: Elaboración propia.

El diseño anterior define fundamentalmente objetivos, metas de resultados y de impactos esperados, estrategias e instancias responsables. Pero sucede que en este proceso pueden generarse al menos tres brechas o diferencias importantes: 1) la brecha de pertinencia del programa y sus objetivos a las necesidades insatisfechas identificadas como problema; 2) la brecha de congruencia entre dichos objetivos, las metas correspondientes y las políticas,

programas y estrategias definidos para alcanzarlos; y 3) la brecha de gestión o implementación del programa que introduce aspectos como la estructura, los procesos y la cultura organizacional de las instancias participantes, así como su sistema de interrelaciones.

Como consecuencia del proceso descrito se obtiene un cierto nivel de resultados e impactos sobre el problema original, que normalmente difiere del planeado debido a la presencia de las tres brechas señaladas, a las que se suma la posible incidencia de variables exógenas tanto en el problema inicial como en el programa.

En consecuencia, puede haber razones endógenas y exógenas que expliquen que un programa no logre sus objetivos en la forma esperada, o que a pesar de lograrlos no contribuya a resolver el problema social que le dio origen.

Para identificarlas se realiza la evaluación de programas. Puede procederse con base en tres modelos diferentes de análisis:

- Evaluar únicamente algún elemento del proceso: sólo los procedimientos empleados en la gestión o evaluación en estudio, o bien, sólo los resultados alcanzados;
- evaluar todo el proceso: diseño e implementación en sus diferentes componentes (insumos, procesos de transformación, resultados e impactos); o
- comenzar por la medición de impactos y, sólo si éstos no son satisfactorios, retroceder a analizar las brechas previas para encontrar las causas del fracaso y diseñar las correcciones posibles.

Debido a que una evaluación de impacto es difícil, costosa y sus resultados no suelen ser contundentes, a menudo se opta por el primer modelo de evaluación, en el que el impacto no se mide rigurosamente, pero se identifican las brechas que, dentro de la lógica del programa, se supone que deben estar disminuyendo la efectividad del mismo, y se procede a su corrección.

Si hubiera buenas razones para pensar que la medición de impacto se puede hacer con suficiente precisión, y motivos para creer que muchos de los programas aplicados son exitosos para la solución de problemas, se podría aplicar a todos el tercer modelo, procediendo a analizar todo el proceso sólo cuando se presenten resultados insatisfactorios, lo cual ahorraría bastante tiempo y dinero. En el desarrollo de este trabajo iremos viendo qué tan realista es pensar que podemos medir rigurosamente los impactos de un programa.

En términos generales, prácticamente todos los autores revisados esperan que la evaluación contribuya a la toma de decisiones y rendición de cuentas a la ciudadanía, coadyuvando a un empleo eficiente de los recursos y a una alta efectividad de su impacto. Algunos autores la consideran un componente crítico de la reforma del sector público (Ospina, 2001). Diversas perspectivas la asocian con el interés público, el mejoramiento de la sociedad y su contribución a la democracia (Scriven, 1967; Weiss, 1982). En la práctica, es necesario analizar en qué condiciones ayuda a realizar juicios y procesos de decisión más razonados y rigurosos (Thoenig, 2002).

### Tipos de evaluación

Existen muchas formas de clasificar a las evaluaciones (y metaevaluaciones), por lo que revisaremos sólo las que nos parecen de mayor utilidad para el presente trabajo. Se aclara que se consideran exclusivamente las evaluaciones realizadas en forma sistemática, objeto de nuestro trabajo, que implican una rigurosa recopilación de datos, basada en un buen conocimiento de epistemología, reglas empíricas, ciencias informativas, proyectos de investigación, valoraciones, estadísticas, comunicación, etcétera, descartando las evaluaciones efectuadas en forma intuitiva, según la clasificación de Stufflebean y Shinkfield (1987).

a) Ya hemos hecho referencia a una primera clasificación en: evaluación *ex ante* o *a priori* y evaluación *ex post* o *a posteriori*, vinculadas respectivamente al “conocimiento en” y al “conocimiento de”, de Laswell. Adicionalmente, hay autores que se refieren al monitoreo o evaluación *ex tempore* o concomitante con el proceso de implementación; se trata en este caso de la utilización de herramientas propias del control y seguimiento de lo planeado. La evaluación *ex ante* se refiere a la comparación de los efectos esperados de las distintas alternativas que pueden permitir la resolución de un problema y constituye una de las etapas de un proceso decisorio que trabaja con base en datos estimados. En el presente trabajo, nos centraremos en la evaluación *ex post*, de aquellas acciones ya ejecutadas, que componen la gestión de los programas públicos seleccionados, por lo que sus efectos constituyen datos históricos.

A pesar de que este tipo de evaluación identifique, mida, compare y califique los resultados o impactos obtenidos, eso no significa que se la deba considerar como la evaluación final propia de una actividad acabada, sino que puede constituirse en un proceso de seguimiento continuo de reflexión

y análisis sobre los avances y limitaciones que se generan en el desarrollo de una actividad, lo que constituye una oportunidad para adoptar medidas correctivas oportunas y permitir un proceso de aprendizaje y crecimiento de las personas involucradas.

b) La evaluación puede ser de carácter externo, en el caso en que es desarrollada por personas u organizaciones ajenas a quienes desarrollan la actividad evaluada. Puede tratarse de una evaluación intergubernamental (Rieper y Toulemonde, 1996), basada en modelos del tipo *top-down*, o independiente, desarrollada por una organización totalmente ajena al gobierno (centros de investigación, organizaciones no gubernamentales, periodistas, partidos políticos, sindicatos, etcétera).

En el caso opuesto al anterior, se habla de una evaluación interna, como la que puede desarrollar el propio organismo responsable de un programa público. Si quienes hacen la evaluación son los mismos funcionarios que aplican el programa, o la población participante en el mismo, se trata de una autoevaluación.

Una evaluación interna garantiza un mayor conocimiento de la actividad revisada y disposición de información, mientras que una evaluación externa asegura mayor objetividad. Sin embargo, conviene resaltar que la evaluación no constituye un proceso libre de intereses, sino que por el contrario, existen personas y grupos preocupados porque los resultados que se obtengan demuestren el éxito de la gestión, mientras los opositores se obstinan en lograr lo contrario. Por esa razón, pueden existir incluso presiones sobre los evaluadores que provocan sesgos en los resultados, disminuyendo la “objetividad” que sería deseable en dicho proceso.

c) Los destinatarios de la evaluación, pueden ser muy diversos. En la evaluación de programas públicos la mayoría de los destinatarios pertenecen a la administración pública, que las realiza para retroalimentar su proceso decisorio; al Congreso, para servir de contrapeso del Ejecutivo; y a la ciudadanía, para fomentar el debate público y la participación ciudadana. Algunos de ellos pueden involucrarse en la realización de la evaluación, que adquiere entonces, el carácter de participativa, misma que, cuando incluye a los ciudadanos interesados en el programa, contribuye a afianzar los procesos de democratización (Blanet *et al.*, 2000).

Esta evaluación, que Guba y Caracelli (1989) denominan “evaluación de cuarta generación” se organiza en torno a las peticiones, demandas, asuntos y cuestiones planteadas por los propios implicados, mismas que constituyen la base para determinar la información necesaria. Su metodología enfatiza

el desarrollo de procesos de aprendizaje de los participantes (Isaac, 1999 y DeBono, 1999).<sup>14</sup>

d) Otra clasificación posible apareció insinuada al tratar el tema del momento en que se evalúa. Se trata de la que identifica los dos tipos siguientes: a) sumatoria, de balance, recapitulativa o conclusiva, y b) formativa, endoformativa o correctora (Scriven, 1967). La primera evalúa los resultados finales de una actividad, plan, política o programa aplicado, en comparación con los objetivos planeados o los resultados obtenidos en empresas similares. Constituye generalmente un instrumento de gestión, dirigido a personas ajenas al desarrollo de la actividad, que sirve para decidir si la misma debe proseguir, con o sin ajustes. La segunda realiza un seguimiento continuo de la actividad desarrollada para detectar a tiempo si se están presentando dificultades y tratar de resolverlas. Cuando se realiza en forma participativa, funciona como un instrumento de aprendizaje colectivo que permite a los mismos protagonistas de la actividad modificar conductas y mejorar acciones, contribuyendo así a los procesos de participación y democratización; en caso contrario, se limita al aprendizaje organizacional del gobierno (Leeuw *et al.*, 2000).

e) El primer tipo de evaluación identificado por Scriven, enfatiza la cuantificación de los resultados hallados, mientras que el segundo hace su análisis en forma fundamentalmente cualitativa, se orienta hacia un conocimiento comprensivo de lo que está sucediendo. Esto nos genera una nueva clasificación, que nos marca diferencias entre una evaluación cuantitativa y cualitativa. Según Hofstède (1980), la evaluación cuantitativa sólo es posible cuando los objetivos perseguidos son claros y los resultados medibles y da lugar a controles de rutina, por expertos, por medio de pruebas e intuitivos. Cuando los objetivos son claros pero los resultados no son medibles, corresponde un control “a criterio” y, en los casos en que incluso la definición de objetivos es ambigua, el control sugerido es de tipo político. Estos últimos dos casos implican la realización de evaluaciones cualitativas, basadas en criterios personales y, por lo tanto, subjetivos.

f) También suele diferenciarse entre evaluación sectorial, enfocada a las instituciones, planes o programas dirigidos a un área sustantiva de política, y evaluación global, que involucra diferentes sectores relacionados con una política.

<sup>14</sup>La participación social tiene un fuerte arraigo en los países europeos, especialmente los nórdicos, pero también se ha desarrollado en contextos muy diferentes. En Japón se le conoce a través del concepto “ba”, que describe un lugar de aprendizaje, de carácter físico, virtual o mental, con el objeto de relacionar individuos u organizaciones, para facilitar el aprendizaje y desarrollo creativo, a través de procesos de socialización, externalización, combinación e internalización (Nonaka y Nishiguchi, 2001). Se trata de una forma de colaboración que, en occidente, podría considerarse como formación de “capital humano” y “capital social”.

g) Stufflebean y Shinkfield (1987) hacen una diferenciación importante entre pseudoevaluación, cuasievaluación y evaluación verdadera.

La primera intenta presentar conclusiones erróneas como fruto de un “evaluación”. Incluye las evaluaciones políticamente orientadas (partidistas, por ejemplo) o “investigaciones encubiertas”, que buscan un resultado positivo o negativo, independientemente de cualquier valoración objetiva. Lo mismo sucede con los estudios basados en las relaciones públicas, cuya intención es ayudar al cliente a lograr una imagen positiva de algo.

La siguiente, da respuesta a algunos asuntos de interés, en lugar de determinar el valor y el mérito de algo, objeto de la evaluación verdadera. Incluye los estudios basados en objetivos (Tyler) y en la experimentación que busca relaciones causales pero no emite juicios de valor.

Las verdaderas evaluaciones se componen de estudios orientados a la toma de decisiones (Cronbach), los centrados en el cliente para valorar sus contribuciones, los estudios que miden los méritos relativos de diversas políticas y programas (Rice) y los que ayudan al consumidor a elegir los bienes y servicios más convenientes.

b) Finalmente, denominamos evaluación integral a la que cubre, utilizando el enfoque de sistemas, los insumos, el proceso, los resultados y los impactos. Dicha integralidad se complementa con el estudio pluridisciplinario de los elementos anteriores. Pero también existe un conjunto de trabajos (Nioche, 1982; Stufflebean y Shinkfield, 1987; Monnier, 1991 y 1995; Mény y Thoenig, 1992) que aportan la identificación de diversos tipos de *evaluaciones parciales*:

- De *objetivos*: analiza si éstos han sido definidos con precisión.
- De *pertinencia*: vincula los objetivos del programa con los problemas que los originan.
- De *coherencia*: relaciona los objetivos con los medios para alcanzarlos.
- De *contexto*: revisa las condiciones en que se ha producido la aplicación de la política o programa así como los procesos de interacción del programa y su entorno.
- De *insumos* o de *medios*: analiza la cantidad y calidad de los recursos empleados y verifica si los mismos han estado disponibles en el lugar y tiempo preestablecidos.
- De *proceso*: revisa los mecanismos que permitieron la transformación de insumos en resultados.
- De *realizaciones, resultados o productos*: intenta apreciar los productos o servicios generados por la actividad evaluada.

- De *eficacia*: verifica si los resultados han permitido lograr los objetivos.
- De *metas*: analiza lo logrado en comparación con lo planeado.
- De *eficiencia*: relaciona los resultados con los esfuerzos realizados para lograrlos (costo-beneficio, costo-eficacia).
- De *efectos*: mide las consecuencias provocadas por los resultados alcanzados en el entorno socio-económico, sin juzgar los objetivos del decisor, incluyendo las externalidades provocadas, en tanto efectos no previstos, positivos o negativos; busca apreciar las modificaciones de la situación o de los comportamientos de los individuos, empresas o comunidades afectadas por la política. Los efectos inesperados se producen debido a cambios del entorno, errores en la estructuración causal de la acción o grandes consecuencias colectivas provocadas por pequeños efectos, y demuestran la ausencia de linealidad causa-efecto de la acción colectiva.
- De *impacto*: se refiere al impacto global, provocado por la combinación de efectos atribuibles al programa, sobre la problemática que justifica la existencia de la actividad (Nioche conjunta en la evaluación de impacto también la incluida aquí como evaluación de efectos).
- De *satisfacción*: determina el grado en que la población considera satisfecha su necesidad o resuelto su problema.

En materia de medición de impacto, si bien el aparato administrativo público está fuertemente diferenciado y cuenta con una débil capacidad de integración, es necesario señalar que la transversalidad de las políticas obliga a realizar evaluaciones que no respetan los cortes institucionales, requiriendo un mínimo de acuerdo y cooperación entre los actores y los niveles de acción y responsabilidad (Duran, 1999).

### Actitudes y enfoques

a) Una actitud que comúnmente se espera de un evaluador es que realice su trabajo en forma tan objetiva como sea posible. Este es un aspecto muy importante dentro de ciertos enfoques metodológicos (positivismo, conductismo), muy cuestionable y poco realista para otros (relativismo) y una falsa dicotomía en el caso de Pierre Bourdieu. En el caso concreto de la evaluación, aun suponiendo que pueda mantenerse un alto nivel de objetividad en la consecución de los cuatro primeros pasos que la conforman (identificación, medición, comparación de efectos y explicación), el quinto implica un juicio de valor que, por mucho que se fundamente en criterios



preestablecidos, no dejará de tener elementos de carácter subjetivo. Lo mismo sucede con el sexto paso, aunque probablemente en menor grado. Adicionalmente, pueden existir durante el proceso de evaluación diversos tipos de compromisos que disminuyan aún más dicha objetividad: afectivos, de asociación, factuales, fiscales, filosóficos y políticos, que deben tratar de evitarse.

En ese sentido, Duran expresa: “Si hasta aquí se ha tratado de consagrar la autonomía de las esferas de los valores y los hechos, la evaluación viene a recordar su interdependencia... La política sigue siendo asunto de juicio y compromiso” (Duran, 1999: 188, traducción libre).

Esta combinación objetiva-subjetiva de la evaluación hace que su credibilidad dependa no sólo de la *expertise* del evaluador, sino también de la confianza que éste genera con base en sus características morales, que le permitan un manejo honesto de la información y una emisión no comprometida de su juicio valorativo; mismo que resulta más garantizado cuando la evaluación queda a cargo de una instancia colectiva y plural, como veremos más adelante. La deontología, basada en los principios de autonomía financiera e institucional del evaluador para garantizar su objetividad, y, de replicabilidad posible de los resultados por otro evaluador que utilice la misma metodología, aparece como un tema central de la evaluación desde el trabajo de Deleau *et al.*, (1986), que la coloca al mismo nivel que los problemas vinculados a su método y forma de organización.

García (en *Diccionario Crítico de Ciencias Sociales*, [www.ucm.es](http://www.ucm.es)) plantea seis razones para destacar el carácter político, no objetivo, de toda evaluación: 1. las políticas y los programas públicos sometidos a la evaluación son aprobados y financiados dentro de un proceso político; 2. la evaluación reproduce los rasgos del contexto político en que surge y sus estructuras, actores y procedimientos influyen fuertemente en su realización; 3. los gobiernos ven en ella una oportunidad de control y un elemento de legitimación, razón por la que la impulsan y financian; 4. en la evaluación misma se expresan intereses políticos contrapuestos, que buscan aumentar sus cuotas de poder e influencia; 5. el diseño de la evaluación y sus criterios son de naturaleza política en la medida en que implican tomar posición sobre la participación de grupos en la misma o la apropiación de la información generada; y 6. se realizan para influir en el proceso de toma de decisiones política, para ganar legitimidad, fundamentar argumentos, etcétera.

De esta forma, García señala que varios estudios han servido para mostrar el mal estado en que se encuentra la evaluación financiada con presupuesto público en términos de credibilidad ante la injerencia creciente del

gobierno en la elaboración de la propuesta y la imposición de restricciones al uso y difusión de resultados, establecidas en el contrato.

b) Mény y Thoenig (1995) identifican cuatro posibles actitudes en un evaluador: descriptiva, clínica, normativa y experimental.

La actitud descriptiva es la que adopta un evaluador que pretende hacer un listado exhaustivo y neutral de los efectos de una intervención, sin emitir ninguna valoración. Más que un evaluador, se trataría de un informante.

En cambio, un evaluador con actitud clínica, complementará la descripción anterior con una explicación relativa al nivel de logro de los objetivos perseguidos por la gestión, basada en el análisis de las condiciones de la ejecución de la misma.

A diferencia del anterior, un evaluador con actitud normativa analiza los resultados encontrados a la luz de sus propios valores, que pueden diferir de los que el programa pretendió alcanzar, entra al debate político y normativo, al tener que tomar partido por alguno de los valores, datos y plazos que existen en la sociedad.

Finalmente, la actitud experimental es la más ambiciosa desde el punto de vista científico ya que pretende establecer rigurosas relaciones de causalidad entre la gestión en estudio y los efectos que se le imputan.

Mény y Thoenig se refieren también a la existencia de evaluaciones como actividad política, donde la información es intencionalmente descartada o tergiversada a los efectos de que se ajuste a los prejuicios culturales, intereses y marcos cognitivos de cada persona; por lo que, en estos casos, la información no constituye la variable esencial para la formulación de los juicios de valor.

c) En materia de enfoques, la evaluación sufre la influencia del debate epistemológico entre corrientes positivistas y constructivistas. Los enfoques positivistas se caracterizan por poner distancia respecto del objeto, preocuparse de la validez interna y externa de su diseño metodológico, la confiabilidad, el poder de generalización, la objetividad, el énfasis en la cuantificación y la utilización privilegiada de la experimentación.

Los construccionistas (R. Mayntz y la Escuela de Colonia, Duran y Thoenig) creen en la existencia de tantas realidades como individuos (en lugar de visiones objetivas), se interesan en fenómenos que articulan lo micro y lo macro, lo formal y lo informal, la organización en red, rechazan la separación tajante de las etapas del proceso de políticas públicas y la realización de estudios de caso sectoriales, y enfatizan la investigación cualitativa, a veces apoyada en la estadística.

Monnier (1989 y 1995) considera que es posible clasificar los enfoques de evaluación en: tradicionales (jurídico y técnico-económico), basados en el positivismo, y nuevos métodos (endoformativos, pluralistas y efectuados por los administrados), fundamentados en el constructivismo.

El modelo jurídico comenzó aplicando el control jerárquico de las prácticas administrativas, al que luego se agregaron los aspectos técnicos y financieros, manteniendo la intención de comprobar, por parte de alguna instancia de la élite administrativa, si la acción realizada se ajustó a los medios y procedimientos definidos por las reglas preestablecidas. No intenta medir la eficacia de la acción, mucho menos su impacto social.

El segundo modelo equipara la gestión pública y la privada. En materia de evaluación, supone la existencia de objetivos explícitos y contextos estables, enfatiza lo cuantificable y excluye todo elemento ajeno a lo técnico-económico. Se desentiende de la dinámica real de los procesos, por lo que genera efectos perversos en entornos de alta incertidumbre. Mejía (2003) le identifica dos limitaciones principales: la irrelevancia de una evaluación que no responde a las necesidades de los actores-clave (gerenciales o sociales) y su limitada institucionalidad, debido a que goza de un débil apoyo gubernamental y no gubernamental.

En cambio, Monnier propone enfocar la evaluación hacia el sistema de actores sociales involucrados, con una actitud pluralista, cuyo antecedente lo constituyen las evaluaciones participativas, denominadas de Cuarta Generación por Guba y Caracelli (1989). Dicha actitud se fundamenta en el reconocimiento de la diversidad de sistemas de valores que coexisten en el seno de la sociedad y se construye a partir de los conflictos sociales emergentes; lo que lleva al evaluador a actuar como conciliador entre las partes en desacuerdo, en búsqueda de una solución satisfactoria, no óptima.

Monnier (1991) plantea, además, que la utilidad social de una evaluación depende, no sólo de la metodología empleada, sino de dimensiones políticas como el acceso a la información, la existencia de condiciones que favorezcan la utilización de sus conclusiones y la credibilidad de la misma para los distintos actores, determinada por la legitimidad del proceso seguido, ya que ninguna evaluación puede considerarse neutral.

Bajo el enfoque pluralista, Monnier afirma que: "...un evaluador experimentado reúne aptitudes, a nuestro juicio, para llevar a cabo una evaluación en un sector que no le resulte familiar. Su desconocimiento le permite salvaguardar la disponibilidad mental indispensable" (Monnier, 1995: 271); en cambio, Mény y Thoenig (1992) consideran que el evaluador debe ser un experto en el tema objeto de su evaluación.

e) En el mismo sentido, Duran (1999) señala que la mayoría de quienes solicitan una evaluación pretenden conocer todos los efectos de una política o programa y poner en evidencia los esquemas de causalidad que los producen. En función de ello el rigor metodológico se ha confundido con un enfoque tecnicista, descuidando la construcción del objeto del estudio y la elaboración de un consecuente protocolo de investigación. Se olvida que las ciencias se limitan a la aplicación de métodos y técnicas y que, aunque la verdad sea única, existen múltiples puntos de vista en un trabajo interdisciplinario y participativo, por lo que los primeros no garantizan totalmente la pertinencia científica de las conclusiones. Debido a que la ciencia no cuenta más que con procedimientos de verificación limitada, es que proliferan más las evaluaciones de proceso que de resultados.

f) Hace ya más de 20 años, Nioche planteaba la conveniencia de avanzar del enfoque de evaluación-sanción al análisis-aprendizaje. Duran y Monnier (1992) también se han declarado a favor del aprendizaje a partir de la experiencia de enfrentar problemas mediante la aplicación de soluciones nuevas, en lugar de intentar exhaustivos ejercicios de programación de objetivos, metas, estrategias, etcétera. Se requiere de una cierta indeterminación, por tanto de libertad, para que se pueda actuar en forma estratégica (Duran, 1999).

Engel y Carlsson (2002), por su parte, plantean que existen tres diferentes perspectivas para impulsar el aprendizaje desde las evaluaciones, a diferentes niveles de la complejidad y participación social: 1. el enfoque tradicional de tomadores de decisiones con participación simbólica para legitimar lo realizado; 2. el aprendizaje organizacional de funcionarios, financiadores y demás actores relevantes; y 3. el impulso al aprendizaje de la sociedad, considerado por el Banco Mundial como un bien público global. En estos casos, el evaluador debe recoger y juzgar evidencias, pero también actuar como facilitador de un proceso de aprendizaje, por lo que, implica un cambio de rol y de las características requeridas para cumplirlo (interdisciplinariedad, habilidades de negociación, didácticas, etcétera.)

Obviamente que la participación y el aprendizaje social no están exentos de riesgos, como algunos casos en que han provocado resultados no deseados, lo que lleva a la existencia de autores que se oponen terminantemente a ella (Cooke y Kothary, 2001).

g) Sulbrandt (1993) clasifica los enfoques de la evaluación en: 1) tradicionales y 2) evaluación global. Considera que, salvo excepciones, la evaluación tradicional no ha satisfecho las expectativas en el caso de los programas sociales, de características muy diferentes a los dirigidos a la inversión econó-

mica, y suele ocultar aspectos desfavorables al programa, a la gerencia o a la institución encargada de implementarlo; en consecuencia, no ha conducido a provocar ningún cambio en el manejo de los programas ni en el aprendizaje de los ciudadanos. Además, sus resultados sólo están disponibles para quienes encargan la evaluación, a parte de los múltiples problemas técnicos que se le critican (imposibilidad de atribuir causalmente los resultados al programa, falta de oportunidad, poca relevancia y utilidad, perspectiva de “arriba a “abajo”, etcétera).

En contraposición a la evaluación antes analizada, Sulbrandt propone una evaluación global, que buscaría constituirse en una instancia de aprendizaje que permitiera rediseñar y mejorar los programas, a través de un trabajo cooperativo, a partir de los errores, omisiones y dificultades encontradas en su gestión.

La propuesta consiste en llenar los huecos dejados por el enfoque tradicional, particularmente: 1. contar con una elaboración participativa de criterios de los diferentes actores sociales involucrados; 2. tomar en cuenta lo sucedido durante la etapa de implementación; 3. identificar el grupo-objetivo, funcionarios encargados y demás actores con sus respectivos intereses en el programa; frente a su total omisión; y 4. comprender la lógica interna del programa y sus bases teóricas (cadenas causales).

El nuevo enfoque utiliza datos objetivos y subjetivos, cuantitativos y cualitativos, mezclando una variedad de métodos de recolección de datos, con objeto de aumentar la comprensión del programa, así como la legitimidad de sus resultados.

Para lograr lo anterior se necesita prestar mucha atención a la capacitación de evaluadores competentes e institucionalizar a la evaluación dentro de la administración pública.

Recientemente, el Banco Mundial (BM, 2001) ha incluido el trabajo denominado “Voices of the poor” en uno de sus reportes, que recupera la participación, en discusiones en pequeños grupos y entrevistas grupales e individuales, de 60,000 personas pobres en 60 países. También el Banco Interamericano de Desarrollo (BID, 2000), la OCDE (2001) y otras agencias de desarrollo impulsan el aprendizaje participativo a través del análisis de los éxitos y fracasos de acciones y programas.

En México, la Secretaría de Desarrollo Social (Sedesol) también aplicó en 2003 una encuesta denominada “Lo que dicen los pobres”, dirigida a captar las percepciones y expectativas de la población que vive en diferentes niveles de pobreza sobre bienestar, vida digna, satisfactores mínimos, etcétera (Székely, 2003).

Como puede verse, los temas de la participación y el aprendizaje social aparecen en la caracterización de diversas actitudes que se oponen a los enfoques tradicionales, constituidos por evaluaciones gubernamentales con el objetivo de retroalimentar la toma de decisiones, sin tomar en cuenta el punto de vista de los actores sociales ni informarles de los resultados hallados. Dan lugar a tres posibles niveles de participación de la sociedad civil en la evaluación:

- evaluaciones similares a las tradicionales, pero que incluyen la aplicación de encuestas con preguntas preestablecidas para conocer algunos datos y opiniones de los actores sociales;
- evaluaciones realizadas en la forma anterior, que agregan el empleo de procedimientos de rendición de cuentas a la ciudadanía de los resultados encontrados (Nioche); y
- evaluaciones realizadas con la participación de los actores sociales (Monnier).

b) Actualmente, en Europa (EES, 2002), se habla de tres enfoques alternativos (a pesar de que no siempre se ven tan independientes) dentro de la evaluación actual: aprendizaje, teorías y pruebas. Se trata de una discusión que no está presente en el contexto latinoamericano. El primero se refiere a los procesos ya comentados de aprendizaje a partir de las evaluaciones pasadas, a nivel gubernamental y de la sociedad civil. El segundo recupera las evaluaciones basadas en la teoría (*theory-based approach*), más allá de la mera aplicación de buenas técnicas y métodos, que no siempre ofrecen resultados claros (Huey Chen, Carol Weiss, etcétera); esta corriente se ha fortalecido con el reciente surgimiento de escuelas “realistas”. Finalmente, el tercero retoma viejas corrientes que buscaron darle un enfoque científico al proceso político, trabajando con base en evidencias logradas en pruebas experimentales rigurosas y normas estrictas (Cochrane, Campbell, etcétera).

Muy brevemente diremos también que no hay un enfoque territorial, sino local, de la evaluación de la acción pública. Se trata de evaluar una mezcla de políticas y/o programas aplicados a un mismo territorio. Su ventaja sería que evita el problema de aislar los efectos de cada política, manteniendo sólo las fronteras de los espacios geográficos en que se aplican. La complicación es que todos los programas pueden adjudicarse los resultados positivos y desviar hacia los otros los negativos, o sea no hay una clara responsabilización de los impactos.

Finalmente, a cada actitud o enfoque corresponden diferentes tipos y métodos de evaluación a emplear, que se revisarán en el siguiente capítulo.

## Contexto político-organizativo de la evaluación

Si bien el problema del método es un asunto clave en materia de evaluación, hay otro aspecto muy importante que enmarca la forma en que se desarrollan sus trabajos: las dimensiones contextuales. Se trata de comprender que las administraciones públicas, en cuyo marco se realizan los procesos de gestión y evaluación, se componen de redes de organizaciones en que los factores políticos, culturales, de estructuración y de funcionamiento constituyen factores relevantes que afectan el desempeño evaluativo. El estudio de dichas condiciones, uno de los tres resaltados<sup>15</sup> en Francia por el Informe Deleau en 1986, constituye el enfoque aportado por la sociología de las organizaciones y la sociología política.

Duran y Monnier (1992) plantean que, efectivamente, la complejización de la acción pública y las dificultades del aparato estatal para justificarse son dos de las principales razones que explican el desarrollo de la evaluación, junto con el avance del análisis de políticas públicas. Se refieren a una crisis del sistema de representación en que la política ha entrado en la “era de la sospecha”, signo del deterioro de la democracia representativa como expresión de preferencias, por lo que los gobiernos tratan de resolver problemas concretos al mismo tiempo que de recuperar el mayor consenso posible.

Pero, legitimidad y eficiencia son difíciles de conciliar. Antes el Estado francés y el interés general concentraban místicamente toda legitimidad, actualmente ésta resulta aleatoria y contingente, y debe ganarse a través de la acción eficaz, de la visibilidad de las consecuencias de las elecciones realizadas. Así la evaluación produce al mismo tiempo conocimiento y justificación (Duran, 1999).

Chanut (2002), quien fuera miembro del Consejo Nacional de Evaluación en Francia, resalta la importancia de la dimensión organizacional en los procesos de evaluación y considera que el énfasis puesto en los temas de metodología la ha mantenido oculta y también ha impedido la capacidad de innovación en la materia. Su trabajo plantea:

1. Las organizaciones se mueven en un mundo de restricciones que afectan a la evaluación: diferencias entre la lógica del conocimiento social

<sup>15</sup> Los otros dos temas eran metodología y deontología (reglas del arte).

y la burocrática dirigida a la acción, incompatibilidad de los tiempos de ambas, falta de acceso a la información, ambigüedad por desaparición de fronteras público-privado, participación de nuevos actores, *expertise* de la oferta de evaluadores, procedimientos impuestos, presupuestos asignados, etcétera.

2. La evaluación no ha conseguido aclarar los problemas y confusiones que disminuyen la eficacia organizacional: pérdida de legitimidad, obsolescencia de los dispositivos e institucionalización de prácticas estereotipadas. La credibilidad de las evaluaciones se funda en la autoridad de los evaluadores y se legitima en el marco de estructuras muy dependientes de la turbulencia política, por lo que la obsolescencia acelerada de los programas provoca una débil capitalización de experiencias. Las formas organizacionales determinan el tipo de evaluación realizada y la naturaleza de los resultados: multiplicación de sus objetivos en función del pluralismo; tendencia a la medición de resultados útiles a la gestión administrativa pero descuidando la evaluación de impacto; énfasis en la implementación de políticas y sus disfuncionamientos, que descuidan la concepción misma de la política; muy pocas experiencias de metaevaluación para la capitalización y apropiación de sus resultados; etcétera.
3. Una mayor atención a los problemas organizacionales aumenta la credibilidad y la eficacia de la evaluación. Propone organizar la evaluación de manera heurística, en lugar de procedimientos y métodos excesivamente formalizados; y pragmática, pasando de constituir una amenaza para los evaluados a transformarse en una fuente de información y aprendizaje útil para la acción, mediante una verdadera restitución y discusión de los resultados por los actores, en lugar de la mera entrega burocrática. Considera que de su adaptación depende la legitimidad y sobrevivencia de la evaluación.

En el mismo sentido Hespel, integrante del Comisariado General del Plan, reconoce que “la consolidación de la evaluación dependerá estrechamente de los medios y competencias concretamente aplicados y de la manera como los actores se apropien de sus enseñanzas” (Hespel, 2002: 51, traducción libre).

Lo anterior no es sencillo puesto que los decisores no actúan con una lógica de consecuencias exógenas, sino de acuerdo a una cultura compartida, compuesta de reglas internas y modos de razonamiento implícitos. Esto provoca que los errores de cálculo y las consecuencias no deseadas sean los elementos constitutivos de la acción pública, no aberraciones marginales. Por



lo que se propone redefinir la evaluación como un proceso de aprendizaje institucional, de transformación de lo que conviene hacer, de las reglas y las identidades. Para institucionalizar realmente esa evaluación, se señala que es necesario algo muy diferente a establecer procedimientos administrativos: crear las condiciones que se enfoquen a cambiar la teoría en actos de los decisores, a modificar los modelos cognitivos y normativos que estructuran su visión del mundo, su representación de los problemas y su apreciación de los instrumentos movilizados para actuar (Thoenig, 2002).

### Comentario final

Más allá de la gran diversidad de definiciones, tipos y enfoques de evaluación revisados es posible identificar dos tendencias fundamentales:

- Una, basada en un enfoque positivista, preocupada por la objetividad del trabajo, que se compone de definiciones de tipo instrumental, interesada fundamentalmente en evaluación de tipo sumatoria y desarrolladas desde un enfoque no participativo.
- Otra, fundamentada en el constructivismo, motivada por la credibilidad y legitimidad del proceso y sus resultados, que aporta definiciones de corte sociológico, y prioriza las aplicaciones de evaluaciones formativas con enfoque participativo e interés en el aprendizaje social.

Cada una de ellas se acompaña también de énfasis diferentes en cuanto a la preocupación y selección de los métodos y técnicas de evaluación, que serán objeto de exposición en el próximo capítulo.

## Capítulo 3

# Métodos y técnicas de evaluación

### Introducción

ES EL MOMENTO de referirnos a los métodos y técnicas empleados en la evaluación *a posteriori*, así como a sus respectivos criterios de decisión. Sin embargo, no se realizará una presentación exhaustiva de todos los métodos, lo que excedería las posibilidades de este trabajo, sino que se mencionarán las principales características, aportes y limitaciones de cada uno de ellos y se ofrecerán referencias bibliográficas donde los interesados podrán profundizar en su conocimiento.

### Debate sobre el método

El tema del método de evaluación es un reflejo de toda la discusión epistemológica generada en torno a las posibilidades y formas de acceder al conocimiento que, en nuestro caso, opone fundamentalmente enfoques positivistas y construccionistas (relativistas), pasando por una perspectiva contingente, que utiliza a cualquiera de los anteriores o incluso una mezcla de ambos, dependiendo del contexto de la evaluación.

Beltrán (1991) plantea que, tanto teórica como prácticamente, no existe unanimidad en lo que se entiende por método científico, en tanto método único, detallado, riguroso y universalmente aceptado; sólo una aceptación general de sus principios: contrastación como evidencia empírica, juego entre teoría y realidad, exclusión de la manipulación y el engaño, recurso a la comunidad científica como árbitro y reconocedor de la verdad, y renuncia a la justificación absoluta de la verdad encontrada. Agrega la inconveniencia de que las ciencias sociales tomen como modelo a las ciencias naturales debido a las peculiaridades de su objeto.

Al hablar de métodos nos referimos a procedimientos de investigación que se desprenden de una postura epistemológica y un marco teórico, en que el evaluador, basado en esta perspectiva y bajo un conjunto de principios y

normas, realiza una serie de operaciones que le permiten seleccionar y coordinar ciertas técnicas para alcanzar sus objetivos. Estas últimas constituyen instrumentos concretos de recolección, análisis, tratamiento, síntesis e interpretación de datos,<sup>16</sup> que aparecen mejor desarrolladas que los primeros, los que requieren de una mayor discusión crítica, poco recogida en la literatura sobre evaluación. Esto sucede incluso entre constructivistas (Cook y Campbell, Cronbach), que ponen el acento en los efectos de contexto, resaltando la pluridimensionalidad de los problemas objeto de análisis y la diversidad de enfoques de los actores implicados en la evaluación.

Desde una actitud pluralista, se insiste en que la metodología no se reduce a la selección de instrumentos, "...implica una reflexión sobre el conjunto de procesos que permiten, en una sucesión de interacciones y decisiones más o menos conflictivas, la producción de la evaluación por los actores sociales, institucionales, políticos, administrativos, etcétera" (Lascoumes y Setbon, 1996: 6, traducción libre), por lo que proponen pasar del método como herramienta al método considerado como proceso de construcción, que puede ser analizado sociológicamente.

También resulta preocupante la tendencia a aplicar las mismas recetas metodológicas en diferentes contextos, obteniendo conclusiones a partir del mero tratamiento de los datos, sin tomar en cuenta una teoría explicativa que permita su interpretación (Thoenig, 2002). Identificar la teoría del cambio social implícita en toda política pública es esencial para el evaluador, para comprender la relación de causalidad que se pretende accionar (Mény y Thoenig, 1992).

La evaluación de políticas públicas consiste en poner a prueba, como si se tratara de una hipótesis, explícita o implícita, si una política o programa de acción determinados han sido o no capaces de provocar los efectos esperados. Esto implica asegurarse de que dichos efectos se han producido, y que la política ha sido su única causa, descartando impactos de otras posibles variables intervinientes.

Corresponde aclarar que los métodos y técnicas de evaluación que revisaremos son utilizados tanto en la evaluación de políticas como en la de programas públicos, en parte por la dificultad práctica de diferenciarlos en muchos casos concretos, y en parte, como señala Weiss (1998), porque han resultado bastante adaptables a cualquiera de ellos.

La forma en que la evaluación se lleva a cabo, depende de las múltiples combinaciones de respuestas a las preguntas formuladas en el capítulo ante-

<sup>16</sup> Adaptado de las notas suministradas por el profesor Gilberto Jiménez, durante sus conferencias sobre Epistemología de las Ciencias Sociales, FCPYS, UNAM, 20 y 23 de enero de 2003.

rior, razón por la que hay un gran número de posibilidades para responder a necesidades diferentes. En particular, la discusión sobre la utilización de métodos cuantitativos o cualitativos nace con los orígenes mismos de las ciencias sociales y las propuestas alternativas de la corriente positivista (Durkheim) e interpretacionista (Weber); sin embargo, hay que resaltar que ambos autores emplearon los dos tipos de métodos según lo requiriera su objeto de estudio, por lo que no debe confundirse el uso de un método con la adhesión a una escuela teórica. “Los debates entre comunidades que se identifican con ciertos métodos expresan aspectos más profundos, no siempre explícitos, porque indirectamente se vinculan con el compromiso con ciertos supuestos subyacentes en los paradigmas o con alguna concepción teórica” (Tarrés, 2001: 40).

A pesar de los esfuerzos de la última década por reconciliar estas posturas, la controversia todavía muestra la presencia de comunidades científicas empeñadas en una especie de dogmatismo metodológico.

A continuación se exponen brevemente los principales métodos que la evaluación de políticas públicas ha tomado de la investigación en ciencias sociales.

### Experimentación y métodos cuantitativos

Como se recordará, entre las posibles actitudes del evaluador se incluyó la denominada experimental, que se concreta en el método que lleva el mismo nombre. Dicho método experimental busca establecer relaciones de causalidad entre el programa y su impacto y toma como referencia el paradigma metodológico desarrollado para las ciencias naturales, y aplicado desde hace 50 años por la epidemiología.<sup>17</sup> El mismo ha tratado de adaptarse al contexto de las ciencias sociales debido al alto nivel de rigurosidad que le permite detentar la exclusividad de verificar relaciones de causa-efecto estables entre ciertas acciones y sus supuestas consecuencias (verificación no aceptada como posible por corrientes epistemológicas como el relativismo). Se ha aplicado más a la evaluación de programas que a la vinculada a políticas globales.

Para poner en práctica un experimento social, se requiere construir un estado contrafactual que estime lo que habría ocurrido de no haberse realizado la intervención en cuestión (Banco Mundial, 2004a: 237). Para ello se necesita diseñar la investigación evaluativa desde antes de dar inicio a las

<sup>17</sup> Consultar Schwartz, “Les Techniques de l'évaluation en épidémiologie”, en Nioche y Poinard (1985) sobre el desarrollo de ensayos terapéuticos y de intervención, mediante la realización de una observación simple, una experimentación sistemática y una con elección al azar.

acciones o programas en consideración. Dicho diseño, implica conformar al azar dos grupos entre los destinatarios de las actividades evaluadas, lo que garantiza estadísticamente su equivalencia. En estos dos grupos, supuestamente similares en sus características, se introduce una sola diferencia: la acción, plan o programa a evaluar que se aplica en uno (grupo-caso o tratamiento) y no en el otro (grupo-control o testigo). Lo anterior permite hacer mediciones iniciales y finales (e intermedias si se desea mayor control) del asunto de interés, en ambos grupos, e identificar como causa de las diferentes evoluciones seguidas por los mismos a la actividad desarrollada sólo en el grupo-caso.<sup>18</sup>

Este plan experimental debe cumplir dos criterios (Mény y Thoenig, 1992): una validez interna, lograda a través de la disminución de la posible influencia de factores extraños en la evaluación registrada; y una validez externa, que permita la generalización de resultados a otras políticas o poblaciones supuestamente homogéneas.

La conformación de grupos equivalentes en ciencias sociales, se ha dificultado notoriamente tanto por problemas técnicos como éticos, razón por la cual se ha pensado en diseñar una metodología más flexible, menos rigurosa pero más realista. Surgen así los cuasiexperimentos en que se mantienen las características generales del diseño anterior pero se acepta que los grupos se constituyan en la forma más parecida posible (por apareamiento, por ejemplo), renunciando a la equivalencia fundada en la selección aleatoria. Obviamente, aparece el riesgo de que variables extrañas, no controladas en el cuasiexperimento, afecten los resultados, induciendo errores. Por esto se recomienda hacer un esfuerzo por analizar y descartar esas posibles influencias antes de concluir la probable existencia de una relación de causalidad entre la acción y el supuesto efecto (Campbell y Stanley, 1966).

A menudo, tampoco el método anterior es factible de ser aplicado, ya sea porque no contamos ni siquiera con grupo-testigo no-equivalente, por lo que no es posible comparar resultados; o porque carecemos de información previa al inicio de la acción a evaluar, por lo que no es posible analizar evoluciones “antes-después”. Se habla entonces de la aplicación de métodos no experimentales del tipo “antes-después sin grupo-control”, “después solamente con grupo-control” o “después solamente sin grupo-control”, en que cada uno resulta menos riguroso que el anterior y sus conclusiones menos precisas (Weiss, 1992).

<sup>18</sup> Un ejemplo de este método lo constituyó el conocido estudio dirigido por Elton Mayo en Hawthorne, EUA, en la década de los años veinte.

Los análisis “antes-después” de un mismo grupo (corte longitudinal) o las comparaciones entre grupos (mediante cortes transversales), se realizan a partir de la generación de indicadores para las variables impactadas, siguiendo la lógica de la investigación aplicada. (Cohen y Franco, 1993) o el Marco Lógico aplicado por los organismos internacionales.

Ante la imposibilidad de medir todos los aspectos deseados con una misma unidad de medida (tipo valor monetario), es necesario realizar un análisis segmentado de las distintas actividades y aspecto del logro. Para ello se utiliza un diseño analítico de indicadores parciales, mismo que se trata de complementar con un indicador integrado, que ofrezca una comprensión sintética del logro, para lo cual se necesita tener claridad respecto de la jerarquía de objetivos que permita la adecuada ponderación de su valor (Abate, 2000).

Los indicadores se definen como la expresión de medida de un atributo (Sanín, 1999), medidas cuantitativas de aspectos del logro, a menudo expresadas como ratios o *ratios*, que permiten contar con señales sobre su desempeño (OCDE, 1995), mediante su comparación con los referentes internos o externos (AECA, 2000). Constituyen instrumentos para comprobar si hay o no un acercamiento progresivo a las metas propuestas y permiten operativizar la medición de las variables, señalando sus grados y variaciones en la forma más objetiva y verificable posible.

Su diseño debe cumplir al menos tres requisitos:

1. “permitir comparaciones sincrónicas, en el espacio”;
2. “permitir comparaciones diacrónicas, en el tiempo, y la elaboración de pronósticos”; y
3. “representar valores sobre los que existe un grado de consenso (o posibilidad de lograrlo) entre las partes implicadas o interesadas en la evaluación” (Aguilar y Ander-Egg, 1998: 64).

También se le exigen varias condiciones generales: independencia, verificabilidad, validez y accesibilidad. A estas se agregan la posibilidad de ser medidos (cuantitativa o cualitativamente), contar con información disponible, ser determinantes, o sea, los que expresen mejor la naturaleza del objeto medido (Sanín, 1999) y contar con relevancia, falta de ambigüedad, pertinencia, objetividad, sensibilidad y precisión (Guinart i Solá, 2003).

Como puede verse, el conjunto de requisitos y condiciones es tan amplio que dificulta el adecuado diseño de los indicadores. Incluso, a los anteriores podríamos agregar dos restricciones opuestas que es necesario conciliar:

1. la integralidad o suficiencia, es decir, la necesidad de que el conjunto de indicadores diseñados dé una cobertura razonable a todas las variables identificadas como claves; y
2. su número limitado para no caer en una construcción que se vuelva inmanejable.

Su combinación dependerá de la calidad y cantidad de datos disponibles, de los recursos asignados a la evaluación y del monitoreo previsto en el diseño del programa.

La construcción de indicadores requiere el conocimiento de los objetivos y actividades de la organización, política o programa evaluado, las fuentes de información disponibles, la periodicidad de la información, sus usuarios potenciales, los estándares que servirán de comparación; entre otros. Dichos estándares pueden establecerse con base en valores históricos previos, metas planeadas o resultados de unidades de análisis similares a las evaluadas (a nivel nacional o internacional).

Los indicadores que miden costos, volumen de productos e incluso impacto son sencillos de diseñar en comparación con los indicadores de calidad, cuyo concepto es difícil de operativizar y más aún de establecer sus estándares. En este sentido se han utilizado cuatro formas de construirlos:

1. selección y seguimiento cuantitativo de algunos parámetros significativos;
2. aplicación de encuestas de satisfacción;
3. normalización de aspectos del nivel de calidad deseado (demoras, procedimientos, etcétera); y
4. compromisos o metas a alcanzar en términos de calidad (internos a la administración o difundidos públicamente) (Abate, 2000).

Además, los indicadores directos no siempre son posibles, por lo que a menudo hay que recurrir a sustitutos, con el riesgo de que la medición resulte afectada por factores extraños, sobre los cuales el programa puede no tener control alguno. También pueden presentar efectos no deseables como inducir sesgos que permitan gestiones que cumplen con los indicadores descuidando los aspectos no medidos (calidad muy a menudo) o el logro global de los objetivos.

En el caso de los países en vías de desarrollo, la OCDE (1995) considera que el trabajo debe focalizarse hacia la selección de indicadores-clave en lugar de intentar una excesiva cobertura de los aspectos del programa, sin tener claridad sobre sus interrelaciones y posibilidades de uso.

Los organismos multilaterales suelen solicitar que los indicadores se apliquen siguiendo la orientación de su marco lógico: objetivo general, objetivos específicos, metas, indicadores, fuentes de información y resultados. CEPAL (Sanín, 1999) y el BID (1997) proponen usar una matriz de cuatro filas (objetivo superior, objetivo específico, productos o componentes y actividades) y cuatro columnas (resumen narrativo de objetivos, indicadores para medirlos, medios de verificación de los anteriores y supuestos de desarrollo). En el caso del BM (2004a), se está promoviendo un esquema gubernamental mediante Administración Basada en Resultados (ABR) para la asignación de recursos presupuestales, que incorpora un sistema integrado de monitoreo y evaluación (M&E). Su marco lógico se compone de indicadores para: 1. monitoreo de desempeño continuo: de entrada, actividades y salida (resultados de corto plazo); y 2. evaluación de impacto (de largo plazo).

Los sistemas ABR han sido adoptados por diversos países de la OCDE (Nueva Zelanda, Reino Unido en los ochenta, Canadá a principios de los noventa, EUA a finales de esa década) y existen algunos esfuerzos por implementarlos en AL (Costa Rica, Uruguay, Colombia, Chile, Honduras, Bolivia). Insisten en que la información contenida en los indicadores debe ser entregada, no sólo a los distintos niveles de gobierno, sino a la sociedad civil en su conjunto.

La información cuantitativa generada en los diseños anteriores, puede analizarse con técnicas estadísticas: análisis de datos (tipologías), análisis de series temporales, análisis multivariado (varianza y covarianza) y estimación econométrica sobre datos longitudinales agregados. Éstas también permiten dar tratamiento cuantitativo a la aplicación de encuestas de opinión, otro de los instrumentos de evaluación de resultados más empleado.

Finalmente, también existen técnicas cuantitativas diseñadas para hacer evaluación *a priori* que a menudo se adaptan para evaluar resultados, sustituyendo lo que antes eran previsiones por los datos históricos realmente obtenidos. Lógicamente, en los casos en que utilizan un criterio optimizador, éste se pierde porque generalmente las posibles alternativas de un proyecto no se ponen en práctica y por lo tanto, se carece de posibilidades aceptables de comparación. En su versión original *a priori* dichos métodos asumen las siguientes características (Cardozo, 1993).

El Análisis Costo-Eficacia permite comparar programas y proyectos alternativos que persiguen un mismo objetivo, seleccionando el que presenta menor costo por unidad de logro (C/E). La eficacia suele ser difícil de cuantificar en forma directa, por lo que se tiende a sustituir su medición por indicadores indirectos, supuestamente correlacionados con los objetivos finales.



La mayor limitación del método es no permitir la comparación de programas con objetivos diferentes o múltiples.

El Análisis Costo-Beneficio pretende resolver el problema anterior homogeneizando los costos y los beneficios de toda índole (financieros, económicos, sociales, distributivos, etcétera) por medio de una unidad de medida común (generalmente las unidades monetarias), actualizando los flujos que se presentan en distintos momentos del tiempo, y comparando posteriormente ya sea las relaciones C/B de cada proyecto para seleccionar la menor, o bien, los valores actuales netos (VAN) para seleccionar el mayor de los positivos.

Este método ha sido criticado por: 1. las dificultades y arbitrariedades en que incurre para la cuantificación de efectos que no tienen un valor de mercado, lo que se realiza mediante la atribución de “precios sombra”; 2. la falta de criterios claros para determinar la tasa de actualización social y, en su caso, los factores de ponderación para los diferentes grupos de ingresos; 3. su tendencia a incluir sólo las variables de cuantificación más sencillas (económicas y financieras) a pesar de su objetivo manifiesto; y 4. su excesiva confianza en que un sofisticado análisis, inaccesible para la mayoría de los interesados y basado en cuantificaciones discutibles, puede conducir a una selección óptima.

Los análisis multicriterios, especialmente el Electra<sup>19</sup> diseñado por los franceses, han reaccionado frente a las críticas presentadas a la técnica anterior. También toman en cuenta objetivos múltiples y aplican valoraciones subjetivas y cualitativas en una escala preestablecida, mediante la participación de los grupos implicados en el problema. Proceden ordenando los programas en función de cada objetivo y negociando los pesos relativos de estos con los interesados o sus representantes. Son de carácter no optimizador, limitándose a buscar la solución más satisfactoria a través de una simulación cuantitativa del regateo social que permite identificar el subconjunto de programas que presentan las mejores calificaciones y que al mismo tiempo carecen de fuertes oposiciones.

Finalmente, los *criterios* utilizados en las mediciones cuantitativas giran fundamentalmente alrededor de la eficacia (logro de objetivos), la eficiencia (logro de resultados al menor costo posible), la efectividad (impacto, incluyendo externalidades positivas y negativas) y la equidad (distribución de los objetivos logrados en relación directa con las necesidades existentes). Recientemente se han agregado otros como sustentabilidad y sostenibilidad de los programas. Algunos autores comienzan a mezclarlos para hablar de efecti-

<sup>19</sup> Electra es la traducción de la sigla francesa Electre, que corresponde a Elimination et Choix Traduisant la Réalité, que traducido al español significa Eliminación y Elección que Traduce la Realidad.

vidad sostenible, eficacia viable, eficacia eficiente, etcétera (Hintze, 2003). Cabe aclarar que cuando se habla de logro de objetivos, y principalmente de costos, no se hace referencia sólo a los de carácter económico, sino también a los sociales y ambientales que permiten una conceptualización adecuada de los criterios anteriores, comúnmente reducidos a la productividad o la rentabilidad financiera.

Es importante señalar que han sido muy pocas las evaluaciones que efectivamente han utilizado el criterio de equidad. Esto puede explicarse, en parte, por problemas metodológicos, y por la tendencia preponderante dentro de la teoría económica de no incluir en su análisis los problemas vinculados a la distribución del ingreso, cuyas consideraciones valorativas quedarían fuera del alcance de la ciencia positiva (Weisbrod, en Layard, 1978). Por otra parte, como señala Rosenbaum (2002), el problema de la desigualdad nunca ha sido preocupación central de los gobiernos; sin embargo, en años recientes, a nivel mundial,<sup>20</sup> y tanto en EUA,<sup>21</sup> como en los países de AL<sup>22</sup> y México,<sup>23</sup> en particular, se incrementaron los niveles de desigualdad en forma considerable, por lo que este criterio debería ser considerado explícitamente en la evaluación, lo que resulta de particular importancia a la investigación en curso. Nos referimos tanto a la medición de la equidad a nivel regional, por grupos sociales, étnicos, de género, etcétera.

A manera de cierre de este apartado, puede decirse que la exposición previa muestra que la medición cuantitativa de resultados e impactos no constituye una forma óptima de identificar la efectividad de las políticas y tomar decisiones sobre su finalización o continuación, con o sin modificaciones. Para Nioche (1982), limitarse al paradigma experimental implica renunciar al aprendizaje de efectos no previstos, objetivos ocultos, teorías implícitas, etcétera, y querer imponerlo como criterio fundamental de cientificidad

<sup>20</sup>Según Naciones Unidas (2001), cerca de tres billones de personas viven con menos de dos dólares diarios, mientras que el 20 por ciento más rico de la población global recibe más del 80 por ciento del ingreso respectivo.

<sup>21</sup>Según datos retomados por Rosenbaum, entre 1979 y 1999, el quintil más bajo de ingreso mantuvo su participación en el ingreso total; el segundo, tercero y cuarto, lo disminuyó sensiblemente, de manera que el quintil más alto incrementó su participación del 38.8 al 50.4 por ciento. En un análisis comparativo entre países, el quintil más bajo en EUA tiene ingresos netos por debajo de los que obtiene el similar en Canadá, Reino Unido, Alemania, Noruega y Suecia; en cambio, el quintil más alto de EUA supera al de los países nombrados, salvo Alemania.

<sup>22</sup>América Latina es la región más desigual del globo: el 5 por ciento más rico recibe el 25 por ciento del ingreso (África 24 por ciento, Asia Central 19 por ciento, Asia del Este 16 por ciento) y el 30 por ciento más pobre, el 8 por ciento (África 10.5 por ciento, Asia Central 12.5 por ciento y Asia del Este 12.5 por ciento), según datos del BID para 1998-1999 (2000).

<sup>23</sup>México presenta uno de los índices de Gini más alto de AL (55.0), sólo menor al de Brasil (59.6), Guatemala (59.1) y Chile (57.9) para 1990.

conduce a pseudoexperimentaciones sin sentido o al empobrecimiento del campo del análisis.

El Centre d'Études sur les Réseaux de Transport et l'Urbanisme (Certu, 2002) comparte la idea de que las evaluaciones exclusivamente cuantitativas resultan reduccionistas, debido a la eliminación de los aspectos no cuantificables y sus interrelaciones, pero reconocen su fecunda contribución para fundamentar juicios de valor, elaborar estudios sintéticos y una batería de indicadores bien definidos, con plena conciencia de sus límites.

Por otra parte, en el análisis de las políticas públicas, los métodos cuantitativos no son siempre usados con la neutralidad que aquí han sido expuestos, sino que a menudo son manipulados para justificar y legitimar decisiones adoptadas con base en otros intereses (pseudoevaluaciones), en cuyo caso el modelo incrementalista de Lindblom (1959) explica mejor estos procesos.

### Métodos cualitativos

Es necesario mencionar que algunos métodos mezclan elementos cuantitativos y cualitativos. Para Lascoumes y Setbon (1996) las soluciones aplicadas para alejarse de las críticas sufridas por el enfoque experimentalista, han tendido a esta combinación, como forma de equilibrar los resultados: informaciones cuantitativas recogidas por mediciones comparativas en muestras contrastadas, con datos cualitativos obtenidos por la observación etnográfica (es el caso del modelo de M.Q. Patton retomado por la General Accounting Office de los EUA).

Un método menos cuantitativo de hacer evaluación que los planteados en el apartado anterior (puesto que también reciben tratamiento estadístico), lo constituye la aplicación de encuestas que recogen información, percepciones y opiniones subjetivas de una muestra representativa de destinatarios de las acciones.

Algo similar sucede en cuanto al grado limitado de cuantificación del método Delphi, que se caracteriza por recoger información de expertos, desconocidos entre sí, que envían sus opiniones por correspondencia, mismas cuya media y desviación estándar se analiza y reenvía, para ver si los expertos desean hacer correcciones a sus propuestas, volviendo a repetirse el procedimiento hasta lograr el nivel de consenso requerido. Monnier (1995) considera que la utilización de este método sólo consigue crear consensos artificiales entre los actores, que no sobreviven a la hora de discutir los resultados de la evaluación.

Otro ejemplo podría constituirlo el benchmarking, retomado de la administración privada por la cultura de la gestión y la medición del rendimiento del sector público, que compara los procesos desempeñados en una organización con el líder en materia de “buenas prácticas” de su ramo. Hace uso de indicadores cualitativos en los casos en que no es posible cuantificar las variables a comparar. Pitarelli y Monnier (2000) consideran que se trata de una nueva herramienta, aplicable a la evaluación operativa y la de impacto, que permite juzgar el éxito de un programa por comparación con otro que sirve de modelo, pero advirtiendo que también puede conducir a decepciones, dependiendo de la importancia que se le atribuya en el proceso de evaluación.

Un último enfoque que resulta fructífero es el diseño de escenarios prospectivos en los que se combinan análisis cuantitativos de tendencias con la consideración de los posibles eventos cualitativos que pudieran introducir cambios en las mismas. Esto permite mostrar a los tomadores de decisiones con más claridad, lo que probablemente sucederá en el futuro si la tendencia se mantiene sin modificaciones y cómo podría afectarse si se produjeran eventos que escapan a nuestro control o si realizamos algún tipo de intervención intencional.

Los métodos estrictamente cualitativos han surgido, en su mayoría, de la crítica a un enfoque cuantitativo, de extracción generalmente positivista, que supone la existencia de rasgos universales y regularidades. Así, las ciencias sociales han pasado del estudio de grandes procesos estructurales (macro), con metodologías cuantitativas, a la construcción social de significados en las comunidades locales, con enfoque cualitativo. Los últimos enfatizan el análisis de las diferencias, el estudio de lo individual (micro), de sus motivos, significados y emociones, de las dimensiones sociales, culturales y políticas en que se desarrollan, así como de la vida de los grupos en sociedad, de sus acciones e interrelaciones.

Sin embargo, las críticas al cuantitativismo y las nuevas propuestas no han bastado para que se constituya en un enfoque alternativo, por lo que aún es necesario seguir trabajando en la fundamentación de su confiabilidad y validez, y en la búsqueda de nuevos paradigmas. Además, dichas críticas exigen un conocimiento profundo de los métodos cuantitativos, y lo más frecuente es la adhesión de ciertos investigadores al enfoque cualitativo por críticas mal fundadas hacia lo cuantitativo, o de plano, un rechazo ideológico o traumático a los “números”.

Un análisis serio y crítico de la potencialidad de los métodos cualitativos y de su utilización en la investigación empírica, sin descartar a los de índole cuantitativa, los define como “...la conjunción de ciertas técnicas de recolec-

ción, modelos analíticos, normalmente inductivos y teorías que privilegian el significado que los actores otorgan a la experiencia” (Tarrés, 2001: 16).

Dichas técnicas de recolección de datos se caracterizan por:

- La relación de involucramiento y empatía del investigador con el instrumento utilizado y con el objeto a estudiar (reflexionando, interviniendo y controlando activamente el proceso), tratando de “observar, escuchar y comprender” al otro.
- Su orientación a la búsqueda de significados, ideas y sentimientos subyacentes o latentes en las descripciones obtenidas de las palabras o de las conductas observadas.
- Su prioridad a la narración y valoración de esa subjetividad, tratando de contextualizar las experiencias estudiadas y de interpretarlas teóricamente. No se preocupan por la representatividad de los casos analizados frente al universo total, por lo que no generalizan sus resultados, salvo con el carácter de hipótesis heurísticas.
- Su entusiasta preocupación por adaptar las concepciones y procedimientos sobre validez y confiabilidad para que el conocimiento generado cualitativamente sea aceptado y útil.

Entre los métodos cualitativos, los más recurridos son: la elaboración de estudios descriptivos de caso, biografías, historias de vida, la observación tradicional y participante, la investigación-acción, las dinámicas de grupo, los grupos de discusión, los paneles de opinión, los debates dirigidos, el análisis de textos, las entrevistas, etcétera.

Aplicados a la evaluación estos métodos permiten introducir el análisis de la motivación, intereses, significados, sentimientos, valores, actitudes, comunicación, conductas e interferencias culturales de los implementadores y beneficiarios de un programa. Intentan arribar a conclusiones sin medir los efectos de las intervenciones o bien complementan una medición previa con un análisis explicativo que reconstruye el proceso de su implementación para comprender los resultados alcanzados. Los tipos más utilizados en evaluación son la entrevista, el análisis de textos y los estudios de caso.

La *entrevista* o “...encuentro hablado entre dos individuos que comporta interacciones tanto verbales como no verbales. No es un encuentro entre dos personas iguales puesto que está basado en una diferenciación de roles entre los dos participantes” (Pope, 1979, citado en Rubio y Varas, 1997). Se la ha definido como una situación construida con el fin específico de que un individuo pueda expresar, al menos en una conversación, sus referencias pasadas

y/o presentes, así como sus anticipaciones o intenciones futuras (Kanh y Cannell, citado en Tarrés, 2001). A diferencia de la encuesta estadística, la entrevista utiliza un muestreo de tipo teórico o intencional y se repite hasta alcanzar un “punto de saturación”, en la que el investigador considera que su continuación no agregaría información relevante. Cuando se dirige a informantes-clave intenta recoger información de aquellos que la poseen en forma privilegiada debido al papel que juegan en el proceso en estudio (directivos de programas, líderes de grupos, etcétera).

La entrevista también puede asumir diversos tipos, de los cuales los más empleados en evaluación son: 1. semiestructurada, combinación útil para entrevistar personas que disponen de poco tiempo (directivos, burócratas, líderes, etcétera); 2. enfocada, en la que el investigador conoce de antemano la participación que ha tenido el entrevistado en algún evento de su interés y asume una posición directiva de la entrevista, conduciéndola hacia el asunto que constituye su foco de atención, pero dejando en libertad al entrevistado en la exposición del mismo; y 3. grupal, dirigida a un grupo focal sobre un tema compartido por todos, donde el investigador actúa como moderador.

El *análisis de textos o del discurso* es un método hermenéutico de análisis estructural, comprensión e interpretación de textos que estudia el contenido de los mismos (hermenéutica interna) y el entorno en que se producen (hermenéutica externa), de manera de “leer” incluso lo que no está escrito expresamente en programas, informes, etcétera. Busca reducir una cantidad importante de datos a una unidad de medida, mediante el empleo de la intuición del investigador. Entre sus tácticas figuran: contar, relevar patrones o temas recurrentes, buscar la verosimilitud o sentido de las palabras, agrupar, construir y utilizar metáforas, separar variables, relevar relaciones entre variables, etcétera (Miles y Huberman, citado en Díaz y Rosales, 2003).

Los *estudios de caso* pueden apoyarse tanto en la tradición cuantitativa como en la cualitativa y complementarse con métodos comparativos y experimentales. El caso se refiere al análisis de algo específico: entidad empírica o construcción teórica que es objeto de la investigación. Analiza el objeto individual a ser estudiado, manteniendo un enfoque holístico, que permite preservar la unidad del todo. Puede ser una forma de comprender mejor el caso particular o de ejemplificar un tema, pero también puede permitir mejorar la comprensión de un fenómeno general. En la elaboración de estudios de caso no se parte de un marco teórico o hipótesis, sino que se privilegia la descripción sobre la estandarización generalizadora, por lo que sólo es

posible pensar en algún tipo de generalización de estos casos a partir de la acumulación de los mismos y la consideración de casos contrastantes, hasta lograr la depuración del modelo. Cuando este estudio recae en una persona, puede tratarse de: una autobiografía, un relato de vida, una historia de vida, o un relato testimonial.

Los *criterios* empleados en estos estudios cualitativos, al aplicarse a la evaluación de programas, enfatizan aspectos vinculados a la calidad, satisfacción del usuario, congruencia, oportunidad y pertinencia, entre otros.

En relación al análisis de la información cualitativa, el investigador trabaja con mucha libertad para dar sentido a los datos en estudio, sin expectativas o teorías preexistentes. Recientemente se han desarrollado programas de computación (ATLAS, NUD\*IST) para el manejo de información proveniente de entrevistas o de apoyo al análisis de textos, que permiten realizar conteos y clasificaciones complejas en poco tiempo.

Lógicamente los planteamientos cualitativos han desembocado también en una redefinición de los criterios de confiabilidad y validez, heredados del positivismo. Mientras los métodos cuantitativos gozan de amplia aceptación en estos términos, entre los cualitativos la preocupación ha aumentado, especialmente en años recientes, pero con un desarrollo lento, insuficiente y diversificado. No resulta fácil lograr un mínimo de rigor científico y concordancia aceptable de la estructura narrativa proporcionada por los informantes a partir de los hechos, tal y como se les presentan de manera cotidiana en su realidad. (Tarrés, 2001). Sherrard y Barrera (1995, en Tarrés, 2001: 89) plantean que la validez y confiabilidad de las entrevistas dependen de tres criterios:

- legitimidad del investigador en el medio académico y ante su informante;
- empleo apropiado de la técnica basado en su experiencia y entrenamiento; y
- calidad del “rapport” (aceptación o entendimiento mutuo) establecido con el entrevistado.

El problema que nos deja este tipo de planteamientos es que resulta muy difícil valorar estos criterios para quienes no participaron en el proceso, y por lo tanto, limitan la confianza para utilizar sus resultados como información secundaria. Sin embargo, hay que reconocer que, tampoco en materia de métodos cuantitativos podemos estar seguros en temas como la composición aleatoria de la muestra para garantizar su representatividad, o la forma en que

se aplicó un cuestionario; por lo que parece que el único criterio que podemos tomar en cuenta, en ambos casos, se refiere a la legitimidad y prestigio del equipo de investigadores y la institución que realizaron el trabajo.

En la evaluación de programas, los métodos cualitativos han sido especialmente utilizados por el enfoque pluralista. Monnier (1995) propone el modelo del *modus vivendi* o de negociación de acción. Es el que considera más adaptable a entornos cambiantes, ya que “La acción será juzgada positiva o negativamente, por una u otra parte, en virtud de que se hayan satisfecho sus intereses específicos en el marco de la acción realizada después de la negociación” (Monnier, 1995: 106).

La legitimidad de este modelo se basa en el compromiso de las partes (gobierno, legisladores, administración pública, asociaciones profesionales, agentes económicos, expertos, investigadores universitarios y público en general), lo que, según el autor, reportaría mayores éxitos en programas públicos innovadores.

Los diversos actores mencionados pueden ubicarse en uno o más de los siguientes contextos: el de legitimación, compuesto por los responsables (jurídicos, científico-técnicos, operativos, etcétera) de la decisión; el de la acción programada, integrado por quienes han recibido un mandato de los miembros del grupo anterior para la ejecución del programa, que disfrutan de cierta autonomía relativa; y el de reacción, conformado por todos los actores sociales que emiten un juicio sobre el valor del programa (evaluación). Esto explica la necesidad de muchas mediaciones para alcanzar un compromiso, siquiera inestable, entre partes tan heterogéneas. Por lo que el autor concluye que: “...cualquier evaluación da lugar a conclusiones relativas; estas conclusiones dependen del punto de vista del actor considerado” (Monnier, 1995: 127-128).

Para Lascoumes y Setbon (1996), mientras que en los estudios experimentales es “el evaluador” quien carga el peso de entregar resultados rigurosos y creíbles (dimensión cognitiva); en la evaluación pluralista, ésta depende de una instancia con amplia participación de actores que representan visiones, valores e intereses diversos, creada específicamente para lograr, además de la dimensión anterior, la producción colectiva que garantice su aceptación como referencia (dimensión política). Se cuidan los aspectos metodológicos y se persigue una doble legitimidad: científica y política.

De esta forma, la evaluación pluralista renuncia a alcanzar conclusiones científicamente verificadas para negociar un compromiso aceptable con los actores. Se trata de un intento colectivo de información y un proceso interactivo e iterativo para la elección de los criterios de la evaluación.



En el contexto francés este enfoque ha generado cierta polémica: Nioche (1992) considera que la propuesta anterior no constituye una evaluación, ya que el objetivo de ésta es informar al mundo exterior (no a los protagonistas, como señala Monnier), por lo que no puede basarse en criterios negociados, no sirve para responder si la política ha sido eficaz o no, si debe continuar o cambiarse. Nioche es partidario de una evaluación externa, que pueda complementarse con la interna, tratando de encontrar su racionalidad *a posteriori*, por una vía no optimizadora.

### Comentario final

Es importante resaltar que en la evaluación no hay métodos, cuantitativos ni cualitativos, que resulten siempre mejores que los demás; el o los métodos adecuados dependerán siempre de la situación a evaluar. Monnier (1995) plantea que estos dependen de la perspectiva de quien la encarga (jurídica, técnico-económica o política), las características del programa (innovador o rutinario), del tipo de evaluación a realizar (de objetivos, de impactos, de proceso o pluralista), de la intención (endoformativa o recapitulativa), etcétera.

A pesar de lo anterior, se perciben énfasis diferentes y muy marcados de las dos tendencias identificadas al finalizar el capítulo anterior: los enfoques positivistas utilizan prioritariamente la experimentación y los métodos cuantitativos, mientras los construccionistas prefieren los estudios no experimentales de tipo cualitativo.

En una revisión hemerográfica realizada hace unos años (Cardozo, 1999), se comparó el papel jugado por los métodos cuantitativos en estas evaluaciones, en las dos fases del proceso de políticas públicas en que éstas resultan más útiles, con los estudios de carácter cualitativo que se publican sobre el tema, concluyendo que esta última perspectiva está siendo más productiva y probablemente, por diferentes razones, más aceptada por los lectores. La tendencia muestra más interés en el desarrollo de los segundos, que se supone predominarán en los próximos años, salvo que se logre mayor realismo en el desarrollo de modelos matemáticos aplicados.

En este trabajo se considera que los métodos cuantitativos constituyen una aportación valiosa al estudio de las políticas y programas públicos, siempre y cuando se tenga presente el conjunto de supuestos en que se fundamentan, y por lo tanto, las limitaciones que presentan, y que se les complementen con el estudio de los aspectos cualitativos, importantes en este campo, como son los problemas de negociación, consenso, legitimidad, satisfacción, cali-

dad, etcétera. De esta manera se obtiene la adecuada integración de las racionalidades técnica y política, necesarias en el campo de las políticas públicas.

En particular, los métodos cuantitativos proveen los instrumentos que hacen posible contar con un panorama global y representativo de los logros y limitaciones de los programas que se evalúan; mismos que resultan necesarios pero no suficientes para la realización de una evaluación integral que pretenda explicar los resultados alcanzados.

Por su parte, los métodos cualitativos se consideran un necesario complemento de los cuantitativos para profundizar en el estudio de situaciones concretas y comprender aspectos subjetivos en mayor profundidad.

Una combinación de ambos tipos de métodos también es recomendada por el Banco Mundial (2004a) para la realización de evaluaciones.

## Capítulo 4

# Experiencias de evaluación en organismos internacionales y países desarrollados

### Introducción

SI BIEN se mencionan antecedentes de evaluación en la antigüedad (China en el siglo XX a.C. o Grecia en el siglo V a.C.), en la época moderna ésta reaparece en el siglo XIX en Inglaterra para la evaluación de servicios públicos y en EUA en el XX en la evaluación del aprendizaje educativo.

A partir de 1980, la mayoría de los países del mundo han intentado aplicarla. En los noventa, con la segunda generación de reformas del Estado, esto ha empezado a incrementarse debido a la necesidad de enfrentar la pérdida de legitimidad del aparato público, a los logros alcanzados por los países de la OCDE en la materia, al auge del NPM, y a la presión de actores sociales, entre otros.

En consecuencia, en el presente capítulo se procede a analizar el papel que han jugado en el tema los organismos multilaterales que actúan en el ámbito internacional y las experiencias de algunos de los países desarrollados con mayor avance en la materia, particularmente el caso francés debido a las similitudes que en su devenir guarda con el caso mexicano (centralismo, resistencia a la evaluación, falta de participación del Poder Legislativo, problemas administrativos, políticos, culturales y metodológicos, entre otros).

Ámbito internacional:  
¿control o evaluación participativa?

Al inicio de los setenta, la evaluación pasó directamente de la burocracia federal de los EUA a algunos organismos de las Naciones Unidas (NU) y a los bancos de fomento multilaterales con sede en Washington (Brown, 2001).

A mediados de los ochenta, la evaluación logra su madurez y deja de tener un enfoque relativamente simple de monitoreo y control, para transformarse en un análisis más sistemático de procesos, orientado al aprendizaje; modelo asumido por la mayoría de las agencias de NU a inicios de los noventa.

Por esas fechas destaca el fomento a la creación de capacidad de evaluación por parte del departamento de evaluación de operaciones del BM, quien intentó aplicar su Sistema de Evaluación de Resultados de la Gestión, brindando asistencia a un banco regional de desarrollo y a seis países, entre los que figura Brasil, Colombia y Chile.

Las agencias de cooperación se habían preocupado por identificar claramente ante la opinión pública los resultados de su asistencia y satisfacer sus propias necesidades de responsabilidad fiscal, excluyendo a los beneficiarios del proceso de evaluación, quienes la habían sufrido como un control impuesto, no como oportunidad de aprendizaje.

Esta situación insatisfactoria fue percibida por parte de los organismos, algunos de los cuales (BID, por ejemplo) modificaron su actitud, abandonando el criterio de “inculpación” que los caracterizó, y sustituyéndolo por el aspecto didáctico de la rendición de cuentas, basada en una cultura de la evaluación que reduciría las posibilidades de corrupción y reforzaría la responsabilidad de los funcionarios públicos.

Hacia finales de los noventa, Spitz (2002) identifica un nuevo énfasis en la evaluación, la medición de logros de corto plazo para la administración basada en resultados, que responde a los imperativos del *value for money* (originado en Reino Unido en el periodo de Margaret Thatcher), con el correspondiente debilitamiento de la orientación hacia el aprendizaje. A pesar de que, en teoría, estos enfoques podrían compatibilizarse, la tendencia entre los administradores de nivel medio ha sido regresar a los mecanismos de monitoreo y control, en el marco de un enfoque de “caja negra” de meras relaciones entre insumos y productos. Spitz considera que las unidades de evaluación (salvo unas 20 dedicadas a desarrollo rural) de las agencias de NU, que desde 1984 han formado parte de la UN Inter-Agency Working Group on Evaluation (IAWG), enfrentan crecientes conflictos de prioridad entre el enfoque cognitivo/de aprendizaje/de construcción de conocimiento y el de medición de logros de corto plazo, en materia de evaluación.

En septiembre de 1999, el BM y el Fondo Monetario Internacional (FMI) anunciaron un nuevo instrumento de política para enfrentar la pobreza en países de bajo nivel de ingresos, como precondition para lograr la cancelación de deudas de los mismos con los organismos mencionados: the Poverty Reduction Strategy Paper (PRSP). Manifestaron mucho énfasis en monitorear el impacto de los PRSP, recuperando el enfoque participativo, tanto de la población como de los distintos ámbitos de gobierno.

Sin embargo, Foresti *et al.* (2002), miembros de prestigiadas organizaciones no gubernamentales (Oxfam, Save the Children), consideran que no

pasó de ser una consulta superficial, ya que un genuino enfoque participativo implicaría abrir espacios para que puedan expresarse las voces que tradicionalmente no se han escuchado. En cambio, los sistemas de monitoreo de los países involucrados en los PRSP suelen estar a cargo de los servicios de estadísticas nacionales, que se concretan a hacer el seguimiento de un núcleo estándar de indicadores de pobreza (inscripción escolar, mortalidad infantil, nivel de ingreso, etcétera), a veces inadecuados para algunos contextos. Proponen su complementación con datos cualitativos que recojan las percepciones de la población en condiciones de pobreza y su evolución, sobre aspectos claves de la misma.

En la actualidad, a pesar de los esfuerzos realizados por medir impactos macroeconómicos y sociales (crecimiento, inversión en capital humano, redes de apoyo, etcétera) de los programas dirigidos a reducir la pobreza, y de los miles de millones de dólares que se destinan anualmente a ayuda para el desarrollo, Baker (2000) plantea que, en organismos como el BM en el que labora, aún se sabe muy poco de cada programa o proyecto específico en términos de eficiencia en el uso de los recursos, impactos en la población y ventajas y desventajas de diseños alternativos.

Las razones señaladas para la resistencia a evaluar, por parte de los gobiernos, instituciones y administradores de proyectos, incluyen altos costos, prolongadas duraciones, complejidad técnica, resultados políticamente delicados, resultados no disponibles en el momento necesario, que no responden las preguntas adecuadas, basados en datos limitados en cantidad y calidad o realizadas con insuficiente rigor científico.

En su opinión la evaluación es especialmente necesaria en los países en desarrollo, donde los recursos son más escasos, y los problemas antes planteados pueden revertirse si se cuenta con el apoyo de las autoridades responsables del programa, una planeación adecuada y oportuna y un presupuesto relativamente pequeño en comparación al costo total del programa o proyecto. Adicionalmente, resalta la importancia de contar con la participación de las partes interesadas desde el diseño de los objetivos y el enfoque de la evaluación, una buena comunicación entre los miembros del equipo durante el proceso y una combinación apropiada de métodos cuantitativos y cualitativos para lograr resultados de alta calidad, creíbles y generalizables. Implícitamente, parece también reconocer problemas en la formación de los evaluadores profesionales, ya que el trabajo que publica es un manual dirigido a ellos, para hacer más accesible la información sobre la evaluación de impacto, en especial la dirigida al desarrollo social.

El BM manifiesta haber “reforzado sus esfuerzos para medir el impacto que tienen las intervenciones sobre los pobres” y estar dedicando varias de sus unidades a su realización, como el Grupo Económico de Investigación para el Desarrollo, departamentos regionales, etcétera ([www.worldbank.org/poverty](http://www.worldbank.org/poverty)). En su trabajo dedicado a la pobreza en México expresa: “Un reto decisivo para una administración pública eficaz es el establecimiento de un sistema de monitoreo y de evaluación (M&E) confiable y eficaz” (Banco Mundial, 2004a: 237).

Esta medición de impacto, que privilegia el largo plazo, también tiene sus opositores (Booth y Lucas, 2002, citado por Foresti), quienes sienten que se descuida la evaluación de los procesos y resultados iniciales e intermedios y la explicitación de los supuestos de los distintos actores, por lo que abogan por un enfoque que integre ambos aspectos, incluyendo la participación de los beneficiarios.

La enorme importancia que los organismos internacionales atribuyen a los procesos de evaluación puede percibirse realizando una rápida incursión en las páginas web de algunos de ellos. También es de destacar la promoción de la evaluación en el marco de la Unión Europea (UE): en 1988 se estableció la obligación de evaluar el empleo de los fondos comunitarios y en 1993 se expidió el reglamento que establece que cada acción comunitaria será objeto de una evaluación *a priori*, concomitante y *a posteriori*. Actualmente la UE solicita evaluaciones externas sobre la cooperación al desarrollo comunitario a través de la Unidad de Evaluaciones de EuropeAid, oficina ejecutora de la programación realizada por la Dirección General de Desarrollo de la Comisión Europea.

Ante el limitado papel real jugado por la población no organizada y los beneficiarios en los procesos de evaluación inducidos por los organismos anteriores, surge la necesidad de hacer referencia a la participación de organizaciones de la sociedad civil, comúnmente conocidas como Organizaciones no Gubernamentales (ONG), y directamente vinculadas a los flujos financieros internacionales de fundaciones y otros organismos donantes.

En la última década, estas organizaciones han reforzado su papel como actores en materia de políticas de desarrollo social y comenzado a valorar la realización de evaluaciones para contar con un mayor conocimiento de las consecuencias de sus intervenciones. También han contribuido a defender el derecho a la información y a operacionalizar la rendición de cuentas en la práctica, al evaluar programas gubernamentales y estrategias de organismos internacionales, como el PRSP, para los que han demostrado contar con algunas ventajas importantes como la disposición de información indepen-

diente de la proporcionada por gobiernos y otros actores de las políticas o su cercanía y sensibilidad para recoger el punto de vista de quienes viven en condiciones de pobreza.

Su participación en estos procesos, aunque numéricamente limitada a la fecha, les ofrece la oportunidad de lograr una mejor comprensión del fenómeno en estudio y de incidir en la toma de decisiones y el cambio en las políticas. Algunos de los factores que afectan la calidad de su participación serían: el grado de desarrollo democrático alcanzado; la relación entre sociedad civil y Estado; las diferentes y a menudo incompatibles expectativas de participación de gobiernos y sociedad civil; y el nivel de representación de las comunidades locales más marginadas.

Obviamente, sus evaluaciones también tienen algunas desventajas, como la desconfianza provocada por consideraciones de falta de capacidad técnica y habilidades específicas, representatividad y legitimidad.

En el marco de la evaluación de los PRSP, en AL, se destaca el papel crítico jugado, en el caso de Bolivia, por el Centro de Estudios de Desarrollo Rural y Agrícola, que monitorea los niveles de pobreza y desarrolla investigación para estudiar su conexión con los PRSP, a efectos de brindar información alternativa a la mostrada a través de los indicadores de estos, y generar un debate en las organizaciones sindicales y sociales, menos involucradas que las ONG (Foresti *et al.*, 2002).

A los limitados avances percibidos en materia de participación de las ONG en la evaluación externa de programas, éstas agregan la creencia de que muchas evidencias son ignoradas y manipuladas, por lo que discuten si la sociedad debe o no continuar dedicándole tiempo y recursos a esta actividad. Otra es la situación en cuanto a la valorización del aprendizaje a partir de evaluaciones realizadas a sus propios programas y proyectos.

### Países desarrollados

Monnier y Barbier (citados en la bibliografía de Políticas Públicas Urbanas, [www.urbanisme.equipement.gouv.fr](http://www.urbanisme.equipement.gouv.fr)) realizaron, a principios de los ochenta, un análisis del estado de la evaluación en los 15 estados miembros de la Comunidad Europea. Encontraron que los mayores avances en materia de publicación de informes de evaluación y en su utilización en el marco de negociaciones políticas entre los poderes Ejecutivo y Legislativo y entre actores nacionales y regionales ocurrían en los países nórdicos, Irlanda, Países Bajos y Reino Unido. En el resto, la ciudadanía y sus representantes aún no sacaban provecho de los trabajos realizados.

Guerrero (1995) presenta un estudio del estado de la evaluación en nueve países desarrollados. Sobresalen por sus condiciones favorables y su nivel de institucionalización EUA, Canadá y Alemania; los Países Bajos muestran condiciones también favorables, pero con un menor desarrollo; Reino Unido, Francia, Noruega y Dinamarca presentan condiciones poco favorables y un desarrollo modesto; finalmente, Suiza cuenta con una institucionalización muy modesta de la evaluación, pero ésta es demandada por la sociedad civil, que además participa en sus procesos. Los resultados de las evaluaciones están totalmente accesibles al público en EUA, Canadá, Francia, Países Bajos y Suiza.

Por su parte, Thoenig (2002) refiere algunos resultados del estudio que hiciera en 1999 en unos 20 países de la OCDE. Considera que hay buenas posibilidades de que se adopten prácticas evaluativas si se respeta su naturaleza pragmática, orientada a la acción, a la generación de conocimiento aplicable y oportuno. Además, plantea que las experiencias más fructíferas serían las relacionadas con evaluaciones concomitantes y participativas, que constituirían un acompañamiento de los cambios en el comportamiento organizacional. Por otra parte, a la ya tradicional presión por la eficiencia, se agregaría actualmente, una menor capacidad del sector público para tratar políticamente los problemas públicos, lo que abriría nuevas oportunidades a la evaluación de la acción pública, que ya no aparece centrada en el Estado, como sucedía con las políticas públicas. El futuro de la evaluación dependerá de su capacidad para adecuarse a los nuevos retos sociopolíticos y coadyuvar a la gobernanza.

La OCDE (2001) y la UE han impulsado reformas presupuestarias en los países miembros en el sentido de abandonar los controles centralizados, manejar su gestión con flexibilidad y privilegiar la obtención de resultados, por los que serán evaluados. Sin embargo, eso no ha hecho que la evaluación sea una realidad generalizada, por lo que en países como Italia la situación es la siguiente:

...la obligación de evaluar los programas financiados por la Unión Europea se traduce en un desarrollo reciente de la evaluación en diferentes dominios de la acción pública. Pero el progreso en la calidad de las evaluaciones y su utilización para fines de aprendizaje organizacional se enfrentan a fuertes resistencias en el seno de la administración y a la falta de competencias metodológicas. Este retraso en materia de evaluación tiene un origen a la vez cultural y organizacional (Stame, 1998: 66).



En este apartado se trata de recuperar algunas experiencias, más avanzadas que las latinoamericanas, ya sea por el énfasis en sus instrumentos metodológicos (EUA) o por las dificultades que han debido superar para el desarrollo de una cultura de evaluación y la implantación de estos procesos en sus países, a menudo muy similares a las que se viven en México hoy. En el contexto europeo, revisaremos el caso del Reino Unido en que su implantación fue más sencilla, y el francés, en que se han enfrentado dificultades importantes. Agregamos el caso canadiense, debido a su paradójico desarrollo.

Hay que resaltar que la institucionalización de la evaluación en estos países ha asumido formas muy diversas. Pollit (1998) nos plantea que mientras en EUA ha seguido enfáticamente el modelo del mercado (es un negocio), en Francia se procedió a su institucionalización jerárquica y en el Reino Unido funciona como un conjunto de redes sectoriales.

#### *EUA: los dilemas de la experimentación*

Stufflebeam y Shinkfield (1987) realizan una útil periodización de la evolución de la evaluación en los EUA:

- 1904 a 1930: se le conoce como el periodo pretylerino, en que se extiende el experimentalismo, iniciado con base en la psicología por Alfred Binet para la evaluación del sector de instrucción pública. En 1912, Starch y Elliot desarrollaron las primeras evaluaciones de los servicios públicos educativos, en materia de enseñanza de las matemáticas, contando con un grupo que había recibido un entrenamiento especial y otro que funcionó como control.
- 1930 a 1945: la época tylerina, en que los análisis de programas de educación y salud pública fueron los más desarrollados, resaltando el diseño realizado por Tyler de un método de evaluación educativa, centrado en el nivel de logro de los objetivos.
- 1946 a 1957: la época de la “inocencia” o la irresponsabilidad social, posterior a la Segunda Guerra Mundial, en que la aplicación de técnicas de investigación social creció fuertemente. A finales de los años cincuenta se había consolidado en EUA y comenzaba a aplicarse en los países más desarrollados de Europa. Así, en EUA, la evaluación se desarrolla en el periodo llamado “sociedad de las reformas” y en Europa durante el “Estado de bienestar”. Los gobiernos fueron los principales demandantes de los procesos de evaluación y la sociedad fue el objeto de estudio que debía ser modificado.

- 1958 a 1972: la época del realismo. EUA utilizó la evaluación para impulsar sus programas de gobierno y controlarlos. Se diseñaron nuevos tests estandarizados a nivel nacional (en contraste con el método de Tyler que enfatizaba la diferencia en la aplicación de objetivos de un lugar a otro) y se utilizó el criterio profesional para valorar las respuestas y verificar periódicamente los resultados. Sin embargo, se empezó a comprender que la evaluación no estaba dando los frutos esperados, lo que se consignó en un histórico artículo de Cronbach (1963), que critica la falta de relevancia y utilidad de las evaluaciones realizadas y recomienda modificar las conceptualizaciones y métodos empleados. En 1965, con el lanzamiento de la Guerra Contra la Pobreza, la evaluación pasó a ser exigida por el Congreso en otros programas sociales, además de los educativos. Posteriormente, se creó el National Study Committee on Evaluation.
- 1973 a la fecha en que escribieron los autores seguidos (1987): la época del profesionalismo, en que el rol de la evaluación se redefinió como herramienta de rendición de cuentas, a través el proceso presupuestal, para medir el desempeño gubernamental en la consecución de objetivos y metas (Boyle y Lemaire, 1998). En este periodo la evaluación fue ganando reconocimiento en los organismos gubernamentales, institucionalizándose en las universidades y formando profesionales especializados (Ziccardi y Saltalamachia, 1997).

A partir de 1968, EUA comenzó a innovar con mucho rigor en el diseño de experimentos sociales que, en opinión de Monnier, fueron desilusionando a políticos y universitarios, por lo que en la actualidad el balance es bastante mediocre. Monnier (1995) identifica otra periodización en su desarrollo:

- De 1968 a 1975: es el periodo de grandes ambiciones innovadoras en manos de universitarios. A inicios de los setenta, el gobierno federal encargaba unas 200 evaluaciones anuales con un costo promedio de 100,000 dólares. Durante el mismo se desarrolló el famoso experimento del “Negative income tax”, con 4,800 familias y un costo de 132.7 millones de dólares de 1983. Posteriormente se llevó a cabo el “Great housing experiment” con 30,000 hogares y un presupuesto de 205.7 millones de dólares de 1983. En el último se hizo un especial esfuerzo por controlar los sesgos que terminaban por aparecer en todos los experimentos y que relativizaban sus resultados. El costo promedio de las

evaluaciones de este periodo fue de 70 millones de dólares y la duración de los experimentos de ocho a 10 años.

- De 1975 a 1983: es la etapa de los avances modestos. Disminuyen los objetivos perseguidos por la evaluación y los medios aplicados, el tiempo se limita a unos tres años y los costos aproximadamente a la mitad. El tamaño de las muestras disminuye poco, pero su selección deja de ser guiada por principios científicos. El trabajo ya no se confía a los universitarios, se contratan agencias públicas y se cuestionan los resultados de las experimentaciones. La competencia entre el gobierno federal y el Congreso por contar con mayor número de evaluaciones llevó a que en 1977 se realizaran 1,315 evaluaciones.
- De 1983 en adelante: se trata de la etapa del reformismo prudente, con mayor retroceso del modelo experimental debido a la poca utilidad de sus resultados. El tamaño de las muestras vuelve a aumentar, pero disminuyendo sensiblemente los costos promedio (ocho millones de dólares). El auge de la evaluación acabó durante el gobierno de Reagan, quien recortó los programas sociales de origen federal.

Actualmente este tipo de estudios es muy poco utilizado en los EUA debido a los problemas presupuestales generados durante las crisis económicas, al regreso de los conservadores que no se motivan con grandes innovaciones y a la desconfianza de políticos, universitarios y opinión pública frente a los costosos experimentos.

Monnier agrega que otra de las causas del fracaso de este enfoque de la evaluación se encuentra en sus postulados erróneos: *a)* racionalidad única y objetivos permanentes; *b)* relaciones lineales y unívocas entre el programa a prueba y sus efectos; *c)* representación utópica de una administración pública imparcial en la aplicación de la ley; y *d)* existencia de una solución única y uniforme, lo que provoca que la experimentación se justifique solamente cuando se pretende la generalización del programa a todo el país, si se comprueba que provoca los resultados deseados.

La Office of Management and Budget (OMB), responsable de la preparación del presupuesto federal, es quien ha coordinado en los últimos años la evaluación de programas gubernamentales. Por su parte, el Poder Legislativo cuenta con el apoyo de cuatro agencias independientes del gobierno, que le permiten actuar como contrapeso: General Accounting Office (GAO), Congressional Budget Office, Office of Technology Assessment y Congressional Research Service. La GAO es, desde 1967, la responsable, a nivel federal, de evaluar políticas y programas públicos. La mayor parte de las solicitudes

de evaluación la realizan los comités del Congreso o sus miembros individualmente; aunque también pueden efectuarse por solicitud del órgano responsable del programa o por obligación legal de la Sunset Legislation, que establece las instituciones, políticas y programas que deben evaluarse periódicamente.

En 1998 se aprobó la Federal Sunset Act que exige la revisión periódica de la eficacia y eficiencia de las agencias federales. Dichas evaluaciones son realizadas por una comisión *ad hoc*, compuesta de ocho congresistas y cuatro particulares, que permanecen seis y tres años, respectivamente, en sus cargos. Trabajan en contacto con la GAO, la OMB, el Comptroller General y los comités del Congreso. Su producto es una recomendación dirigida al último sobre la conveniencia de continuidad o finalización de las agencias evaluadas.

Las condiciones sociales para el desarrollo de la evaluación son óptimas, en la medida en que la oposición cuestiona al gobierno desde el Congreso y ambas partes requieren evaluaciones para fundamentar sus posiciones. Además, existe un número importante de expertos en evaluación y recursos para desarrollarlas.

*Reino Unido:*

*un efectivo trabajo de convencimiento y capacitación*

Si bien desde el siglo XIX existen antecedentes de evaluación de servicios públicos, Saint-Clair (1993),<sup>24</sup> describe los primeros pasos que el Reino Unido dio para implantar la evaluación de políticas. Comenzó por crear un grupo mínimo (dos personas), con recursos suficientes para contratar consultores externos de corta duración; luego, definió los objetivos de la evaluación que se llevaría a cabo (mejorar la gestión, la toma de decisiones, la asignación de recursos, etcétera). Cubiertos estos dos aspectos, comenzó la tarea que cubrió varias etapas:

- a) Se hizo un diagnóstico del estado de la evaluación ya que, al igual que en México, todos los ministerios afirmaban hacerla, pero no pasaban de medir la eficacia de los servicios. Encontraron que los sectores más avanzados en la materia eran los vinculados al mercado: comercio, industria, empleo, etcétera. En materia social iban a la vanguardia salud y educación, pero sólo con evaluaciones de aspectos parciales.

<sup>24</sup>Director adjunto de la Joint Management Unit.

- b) Se propuso cambiar la concepción de la evaluación: de la auditoría y la inspección, a un elemento constitutivo de la gestión administrativa. A diferencia de otros gobiernos que ubicaron a la evaluación en los Ministerios de Finanzas o las efectuaron por medio de despachos de auditores externos, se resolvió la aplicación de evaluaciones internas a los departamentos.
- c) Se diseñó un plan estratégico, enfocando el trabajo a los 300 funcionarios con rangos de subsecretarios y responsables de los grupos de evaluación de políticas para convencerlos de la posibilidad y la utilidad de la evaluación. Éste se desarrolló también en diversas fases:
- Durante tres años, la evaluación se introdujo en el programa de formación de altos funcionarios (jóvenes de 30-40 años), con una duración de seis semanas, a tiempo completo, y organizó en cada ministerio una serie de 50 seminarios de media jornada para todos los altos responsables de la función pública. La misma incluía una exposición de una hora y luego un trabajo en equipo para intentar evaluar una política de su propio ministerio. Finalmente, se hacía un balance para analizar lo que podría mejorarse en materia de organización de la evaluación, diseño de indicadores-clave, etcétera. Contrario a lo que se temía, las respuestas a este proceso fueron muy positivas.
  - Posteriormente, se invitó a los departamentos a proponer uno o dos estudios de caso en áreas que no habían sido sistemáticamente evaluadas, con el objetivo de aplicar los métodos y técnicas aprendidos. Se prepararon documentos de aproximadamente 10 páginas, mostrando los beneficios de la evaluación y las posibilidades de su aplicación en otros departamentos. En paralelo, el Ministerio de Finanzas comenzó a exigir la evaluación como requisito para la aprobación de programas con gasto público.
  - Se dejó a los ministerios la decisión de la importancia que le atribuirían a la evaluación: los más avanzados crearon servicios especializados; los menos, la dejaron en manos de los mismos jefes del servicio, con el apoyo de una célula de expertos en metodología.
  - En 1988 se publicó un manual, "Policy Evaluation. A guide for managers", que se envió a los altos funcionarios, para el desarrollo de evaluaciones de políticas públicas. Sus principios eran presentados en forma muy simple, acompañados de un glosario con los principales términos especializados, a fin de desmitificar la evaluación y evitar los temores que la acompañan.

- d) En 1991, se hizo una nueva revisión de avances: algunos sectores habían integrado la evaluación a sus procesos; otros presentaban avances mínimos.
- e) En 1992, el Ministerio de Finanzas reforzó su importancia, al vincularla más al sistema de asignación de recursos. En paralelo, trabajó con la Comisión Parlamentaria de Control de las Cuentas Públicas para mejorar la presentación de la información al Legislativo.

En resumen, el gobierno británico logró avances importantes en el campo en 1980-1990, en que Margaret Thatcher impuso la evaluación de la eficiencia por medio del concepto de *value for money*, que incluye auditorías de desempeño, exámenes especiales y evaluación de programas a nivel microeconómico. En este enfoque el máximo evaluador es el mercado, compuesto por oferentes y demandantes; de ahí la tendencia a simularlo cuando no existe. Sus dificultades no fueron de orden metodológico, sino vinculadas a la resistencia de los responsables de programas para dedicar tiempo a esta actividad, en detrimento de otras consideradas más urgentes.

La Audit Comision, organismo independiente que evalúa en forma externa la eficiencia de los servicios que brindan las comunidades locales, implantó desde 1994 la aplicación de indicadores no financieros, los que deben publicarse anualmente y han provocado un importante interés entre los ciudadanos. Permiten realizar el *benchmarking* entre entidades similares y analizar la evolución histórica de los mismos.

En 1997, el gobierno británico adoptó un programa nacional piloto, Best Value, que combina el enfoque racional-objetivo de la evaluación de resultados e impactos con otro de tipo argumentativo-subjetivo, que enfatiza el aprendizaje organizacional. Esto debido a que la evaluación racional produce bancos de datos pero no genera información útil o cognición, no provoca cambios de comportamiento en la acción (Thoenig, 2002).

El Tesoro, encargado de la política presupuestaria y financiera, es quien incita a las administraciones a evaluar sus programas y difunde normas metodológicas para su desarrollo.

En el parlamento, el Public Accounts Committee cuenta con el apoyo de la National Audit Office (NAO) y tiene atribuciones para pedir cuentas a los responsables de programas, pero dispone de pocos recursos para cumplir esa función. La NAO está facultada para examinar la actividad de los departamentos ministeriales y de los órganos públicos desde el punto de vista económico, de eficiencia y eficacia. Cada año presenta alrededor de 50 informes del tipo *value for money* al Parlamento (Goldsworthy, 1998).

No existe ninguna autoridad dominante que controle el conjunto de actividades de evaluación, si bien, la influencia de las instituciones centrales de gobierno es importante. Tampoco hay un grupo de grandes firmas que concentre los recursos de evaluación, como en los EUA (Pollit, 1998). Cuenta con una reducida disponibilidad de especialistas en evaluación y poca aplicación de las ciencias sociales a este campo, donde el componente político queda casi excluido. Sin embargo, se ha desarrollado una evaluación crítica en los centros de investigación que luchan por conservar su independencia, que pone en evidencia el carácter discriminativo de los dispositivos oficiales, la reproducción y profundización de las diferencias entre los jóvenes, los efectos perversos de sus criterios en la asignación de recursos, etcétera (Freyssinet, 1998).

En un seminario realizado en 1997 en Francia se describía a la evaluación en Gran Bretaña como más desarrollada que en el país anfitrión en todos los niveles de la administración pública, disponiendo de recursos financieros y humanos más importantes y haciendo un uso más sistemático de sus resultados en la gestión. La misma es realizada mediante la conformación de redes de expertos sectoriales, que incluyen universitarios, profesionistas y miembros de la administración pública, y no se limita a los estudios de *value for money*, utilizando una pluralidad de métodos cuantitativos y cualitativos, incluyendo el empleo de protocolos experimentales, aunque en forma menos frecuente que en los EUA.

Canadá:

*¿evaluación real o mero ritual?*

La evaluación de programas ha sido siempre monopolio del Poder Ejecutivo. Dio inicio en 1978 a cargo de la Oficina del Contralor General del Canadá, aunque existían antecedentes desde los sesenta. Éste definía los estándares metodológicos, formaba parte del equipo que seleccionaba a quienes harían las evaluaciones, dirigía la formación de evaluadores, elaboraba guías y demás documentos de apoyo, etcétera. Su reinado en la materia duró hasta su desaparición en 1995.

En 1987, Canadá era el país más avanzado en materia de evaluación (Saint-Clair, 1993). Cada ministerio contaba con una unidad de evaluación, responsable ante el ministro del ramo, que aplicaba evaluaciones a todo el conjunto, a intervalos regulares, de forma que todos fueran evaluados en ciclos de cinco años. Para ello se realizaban entre 300 y 400 evaluaciones anuales.

Sin embargo, Lacasse (1993) tenía la hipótesis de que muy pocas de ellas se enfocaban a analizar sus aspectos más trascendentes: la pertinencia de la política, la necesidad de su permanencia o las propuestas de cambios, etcétera. Un examen de los informes demostró que se dedicaban sólo unas pocas frases rituales al tema central y la propuesta de algunas “pistas a explorar”. En teoría, los evaluadores sabían que debían revisar la razón de ser de los programas, pero todos consideraban que en la práctica era imposible realizarlo con rigor, inútil intentar convencer a los decisores de su importancia, o incluso ilegítimo, debido a las restricciones que pesaban sobre la actuación del ministerio. También afirmaban que al evaluador no correspondía realizar las propuestas, sino a los consejeros en gestión; además de que nadie se las solicitaba.

Se encontró que las “verdaderas” evaluaciones solicitadas por el contralor no se llevaban a cabo, siendo sustituidas por evaluaciones “realistas”, más débiles, demandadas por los propios ministerios. Además, el evaluador debía cuidar que su trabajo no excediera los límites del propio ministerio y evitar la formulación de propuestas de cambios estructurales que implicaran altos costos potenciales a su cliente.

Se llegó así a una ritualización de la evaluación que no podía cuestionar la existencia misma del programa a cargo del ministerio que solicitaba su evaluación, sólo podía proponer mejoras marginales. Se trataba, entonces, de una paradoja: el país contaba con una amplia institucionalización de la evaluación, pero la forma en que realmente se aplicaba, defendía los intereses ministeriales y del evaluador, quien optaba por colaborar en lugar de cumplir su papel de agente de cambio. Complementariamente, había muy pocos trabajos sobre la utilización de los resultados de la evaluación.

Lacasse concluyó que, a pesar de todo lo escrito en materia de metodología de la evaluación, ésta seguía siendo un instrumento manipulable, por lo que se requería de una instancia capaz de funcionar como contrapeso de los intereses sectoriales y organizacionales. Por otro lado, consideraba imposible intentar evaluar sistemáticamente todas las políticas y hacerlo en profundidad (dilema entre cantidad y calidad), tanto por falta de recursos como por imposibilidad de que el sistema político-burocrático digiriera sus resultados, por lo que propuso contar con algunas evaluaciones efectuadas a fondo, con fuertes dosis de análisis político, aprovechamiento de oportunidades, etcétera, sólo en los casos en que se requiriera de cambios sustantivos y se estuviera dispuesto a realizarlos; mientras el resto continuaba con el mero ritual.

En julio de 1994 comenzaron a aplicarse tres nuevas políticas, entre las que sobresale el examen de los programas, tendiente a una gestión por resultados que pretende saber: 1. si la política o el programa está de



acuerdo con las prioridades del ministerio y responde a las necesidades del momento; 2. si logra sus objetivos de acuerdo a su presupuesto y sin imponer repercusiones no deseadas; y 3. si utiliza los medios más apropiados y eficientes para alcanzar sus objetivos. En el marco de esta política principal se aplican también las políticas de evaluación y de verificación interna. Dentro de la primera se prevé el nombramiento en cada ministerio de un gestor superior, responsable de realizar las evaluaciones, independiente de las actividades evaluadas.

A este examen de programas se atribuye la disminución del gasto en un 12 por ciento entre 1993-1994 y 1998-1999, y su concentración en gasto social. Sin embargo, su éxito para modificar la tendencia previa es sólo parcial, persistiendo múltiples problemas, entre los que sobresalen un 60 por ciento de las evaluaciones sin medición de impactos, debido al énfasis exclusivo en la satisfacción del cliente, o, la ausencia de análisis de su pertinencia, por concentrarse sólo en su proceso de implementación (Turgeon, 1998).

Este autor mencionaba que la evaluación adquiriría tanto la forma *prêt à porter* (positivista) en el caso de programas limitados, de un solo objetivo y experimentales, como su confección “hecha a la medida” (constructivista) en el caso de las políticas interministeriales; sin que ninguna intentara imponerse sobre la otra forma. Desde 1994 se ha tratado de involucrar a los grupos de interés.

El caso del programa de acción comunitaria para los niños tuvo la particularidad de conjuntar los dos enfoques anteriores para dar una respuesta más amplia a sus requerimientos: responder con seguridad a la generalización estadística de sus posibles impactos y adaptar las evaluaciones a las preocupaciones cualitativas de quienes gestionan los proyectos, sobre todo a nivel regional.

Turgeon planteaba que la tendencia era que la evaluación continuara en manos del Ejecutivo (ante un Legislativo absolutamente ausente) y que, si no se introducían cambios, se continuaría sin conocer realmente el impacto de los programas, como había sucedido en los últimos 20 años.

Actualmente, Canadá orienta su gestión por resultados, reemplazando controles *a priori* por la evaluación *a posteriori*, con base en la Ley sobre la Administración Pública del año 2000. Este enfoque se acompaña de una amplia participación de los funcionarios centrada en el aprendizaje, de la utilización de los resultados en la toma de decisiones, de la elaboración de un balance de la utilidad de la evaluación realizada, de una rendición de cuentas obligatoria a la ciudadanía y de un reforzamiento del rol de la asamblea general en la evaluación de la gestión.

Los aspectos considerados en la evaluación son: pertinencia de las políticas y programas al momento de su evaluación; coherencia entre actividades, resultados y misión; nivel de logro de los objetivos; impactos positivos y negativos; y, si corresponde, los medios cómo podrían alcanzarse los objetivos en forma más eficaz, mediante ajustes al programa actual o con la introducción de otro nuevo. Los criterios empleados en su evaluación son, entonces, la pertinencia, la eficacia, la eficiencia y el impacto. Las cualidades que deben integrarse en la evaluación incluyen: la validez, la fidelidad a la situación, la credibilidad y la utilidad.

Cada ministerio u organismo decide qué programas serán evaluados y mediante qué procesos, salvo que se haga por iniciativa del Consejo del Tesoro. También pueden realizarse a nivel interministerial y gubernamental en su conjunto.

Las organizaciones que fomentan la evaluación de programas y la rendición de cuentas son el Parlamento, el Ministerio de Finanzas y el Consejo del Tesoro. Existen muchos especialistas dedicados a la evaluación, provenientes de las ciencias sociales.

*Francia:*

*en búsqueda de un modelo adaptado a su idiosincrasia*

La evaluación aparece en Francia, en los años ochenta, con la expectativa de mejorar la gestión pública y aumentar la democratización de la toma de decisiones. El antecedente inmediato de la evaluación en este país se encuentra en los setenta, en que el Ministerio de Finanzas promovió la evaluación *a priori*, mediante la RCB, similar al PPBS de los EUA.

El periodo 1980-1985 se identifica como de reflexión sobre el tema. Nioche (1982) identificaba problemas sociopolíticos, administrativos y metodológicos que dificultaban su avance. Este autor señalaba, en 1983, al hacer un balance de la situación, que Francia realizaba muchas evaluaciones por demanda institucional, pero estaba muy atrasado en cuanto a la demanda social y ciudadana, tal cual se presentaba en Suecia, Dinamarca o Canadá. En 1983, se llevó a cabo un importante Coloquio Internacional sobre la Evaluación de las Políticas Públicas, que sirvió para revisar los avances realizados y dejó planteadas muchas inquietudes.

En cambio, en el siguiente decenio, se incrementa notablemente el desarrollo de la evaluación. En 1986, se produce el Informe Deleau, realizado con Nioche, Pens y Poincard, a solicitud del gobierno socialista (Deleau *et al.*, 1986). En él se define a la evaluación de políticas públicas

como “identificar y medir sus efectos propios”, con dos características principales: *a)* el énfasis en los efectos de la acción pública hacia el exterior, a diferencia del control tradicional enfocado en los procesos internos; y *b)* su voluntad de medición cuantitativa, sumaria, que evitara que su desarrollo tardara periodos muy extensos. Ésta se basa en dos principios deontológicos: la autonomía financiera e institucional del evaluador que garantiza su objetividad, y la replicabilidad posible de los resultados por otro evaluador que utilice la misma metodología. Se propone un proceso a cargo de alguna instancia habilitada al efecto (gobierno, parlamento, jueces, etcétera), sin participación ciudadana. Monnier (1995) define esta evaluación como de carácter “gerencial”, preocupada por la eficiencia y la optimización de los recursos presupuestales y olvidada de la dimensión política.

En 1988 se discuten dos alternativas: crear un consejo nacional de evaluación, de carácter independiente, o bien un consejo interministerial de evaluación, de tipo gubernamental.

En 1989, se presenta el Informe Viveret al gobierno de Francois Mitterrand, que enfatiza el control democrático. En él, la evaluación de una acción pública se transforma en “emitir un juicio sobre el valor de esta acción”, que sólo los tribunales y los representantes electos tienen legitimidad para realizar, con excepción de los casos de autoevaluaciones. Dicho juicio debía fundamentarse en cuatro garantías para que la evaluación cumpliera con su responsabilidad democrática de movilización de la capacidad crítica de la ciudadanía: independencia de criterio (salvo en las autoevaluaciones), competencia y rigor, transparencia en el enunciado de los métodos y criterios aplicados, así como en los resultados obtenidos, y pluralismo de opiniones en lugar de la persecución de una verdad única. A esto se agrega, salvo casos excepcionales, la regla de publicidad de los resultados. Se trata de una evaluación destinada a enfrentar “la debilidad del Poder Legislativo”, “el monopolio de la pericia en manos de la Administración Central” y “el déficit de las formas de organización democráticas” (Viveret, 1989; citado por Monnier, 1995).

La intención de esta evaluación era también que el decisor actuara con conocimiento de las consecuencias de sus actos, o sea que, buscara combinar su aportación al reforzamiento de la democracia y a la eficacia administrativa. Monnier (1995) la define como una evaluación “democrática”, por contraste con la sostenida en el Informe Deleau y con la evaluación “pluralista”, impulsada por él mismo durante el debate gubernamental previo a la presentación del Informe Viveret. Como consecuencia de este informe, se optó por la creación del Comité Interministerial de Evaluación (Cime), acompañado de la integración del Fondo Nacional para el Desarrollo de la Evaluación (FNDE).

El Cime está compuesto de cinco miembros permanentes y de los ministros implicados en el tópico evaluado. Los primeros son propuestos por el Consejo de Estado, la Corte de Cuentas, el Mediador Nacional, o el Consejo Económico y Social. Este último es un cuerpo consultivo que representa a las organizaciones nacionales socioprofesionales, que puede sugerir el límite de ciertos financiamientos (OCDE, 1995). El Cime decide sobre el inicio de la evaluación, aprueba su metodología, acuerda su financiamiento y recibe el informe. El FNDE cubre el 50 por ciento de los costos de evaluación, mientras el resto lo pagan los ministerios.

En menos de dos años se crearon unas siete instancias permanentes, lo que hace a Duran y Monnier (1992) considerar que Francia entró al tema de la evaluación con mucho formalismo y una cierta precipitación. En la práctica, sólo el gobierno aplicó evaluaciones, bajo la presión de los acuerdos comunitarios de Bruselas y en el contexto de las reformas de centro-izquierda en la administración pública.

En 1990, el gobierno de Michel Rocard emitió el decreto núm. 90-82 que creó formalmente el sistema nacional de evaluación de políticas públicas, basado en principios de transparencia, pluralismo, independencia y rigor, cuya definición fue: “Investigar si los medios jurídicos, administrativos y financieros puestos en práctica permiten producir los efectos esperados por esta política y alcanzar los objetivos que se han señalado” (traducción libre). Como puede verse, se retrocedió a poner énfasis en la evaluación *a priori*, olvidándose del pluralismo y la participación del Parlamento. La evaluación quedó a cargo del Cime, el FNDE y de un Consejo Científico de Evaluación (CCE) que garantizaría el pluralismo metodológico, la independencia de las evaluaciones y jugaría un rol pedagógico importante. Sus miembros, 11 personalidades independientes, con distintas experiencias profesionales y supuestamente especialistas, serían designados por seis años por el Presidente de la República y presididos por el primer ministro.

Una vez que el Cime, organismo con mayor poder que el CCE, decidía las políticas que habrían de evaluarse, se integraba una instancia específica, compuesta de 10 a 12 personas, entre funcionarios, especialistas, investigadores universitarios y responsables de programas, que diseñaban la evaluación (sus métodos, técnicas, fuentes de información, etcétera). El CCE debía emitir un comentario inicial sobre la calidad de la propuesta y finalmente sobre la evaluación realizada, misma que se publicaba junto con ella. También debía hacer un reporte anual sobre las actividades de evaluación desarrolladas y los avances en materia de formación de evaluadores. En realidad, ninguna de estas actividades se cumplió en los primeros años de su creación.

En este periodo proliferaron las instancias de evaluación o “comités de sabios”, según Monnier (1995), cuya independencia y neutralidad garantizarían la legitimidad de la acción del Estado. Además de las ya nombradas a nivel intersectorial nacional (Cime, CCE y FNDE), se crearon otras de carácter permanente: una instancia de política sectorial (Comité Nacional de Evaluación de la Política de la Ciudad), otra vinculada a la aplicación de una ley (Comisión Nacional de Evaluación del Salario Mínimo Interprofesional), y dos relacionadas con programas públicos (Comité Nacional de Evaluación de la Investigación y Comité de Evaluación y Prospectiva de Nuevas Calificaciones Profesionales). En la práctica, estas estructuras no se componían sólo de expertos, sino que integraban miembros del Consejo de Estado y el Tribunal de Cuentas, lo que las mantenía fuertemente vinculadas al tradicional control administrativo y con riesgo de ser absorbidas por el sistema que se proponían modificar, sobre todo porque el parlamento quedaba totalmente al margen y no había desarrollado su especialidad en materia de evaluación (Muller, 1994).

También fue publicada una *Guía de evaluación*, que sistematizaba la experiencia de la Misión de Evaluación y Prospectiva, del Ministerio de Equipamiento. Además, entre 1988 y 1993, se llevaron a cabo cinco seminarios para discutir temas como la relación de la evaluación con la administración, la investigación, la decisión y el juicio valorativo, así como, el rol del evaluador. Por su parte, el CCE definió los cinco criterios que guían la evaluación de políticas en Francia: pertinencia de los objetivos a los problemas que el programa intenta resolver, coherencia entre su concepción y su implementación, eficacia en el logro de sus objetivos, eficiencia en la relación costo-beneficio, efectividad o impacto de las consecuencias globales de la política. Para Crozier (1989), siguió siendo una evaluación tecnocrática, sin participación de los más competentes, los usuarios y administrados, cuyo grado de satisfacción se ignora.

Alrededor de 1995 se produce la etapa de maduración de la evaluación, de balance de lo realizado y también de cierto desinterés por ella. En la práctica se habían presentado sólo 27 proyectos de evaluación,<sup>25</sup> de los cuales sólo 10 se finalizaron, tardando para ello entre dos años y medio y cuatro años. Ninguna de dichas evaluaciones fue examinada en su fase terminal y se reconoció oficialmente que no fueron utilizadas para retroalimentar la toma de decisiones gubernamentales. Las razones de lo anterior fueron: técnicas, porque se encontraron problemas vinculados a la definición de los objetivos

<sup>25</sup> Fueron 10 evaluaciones en 1991, siete en 1992, seis en 1993, cuatro en 1994 y ninguna en 1995.

de la evaluación, su carácter meramente descriptivo, dispersión, heterogeneidad y pérdida de datos, entre otros; y políticas, ya que se enfrentaron problemas de credibilidad en la dinámica del trabajo colectivo y las negociaciones realizadas.

Los sectores que más avanzaron en la evaluación de sus políticas fueron: desarrollo regional y urbano, equipamiento, salud y educación. Monnier (1995) critica el tipo de evaluación gerencial y “apolítica” que se ha aplicado en la educación universitaria, la investigación y la salud, donde más que entre expertos en evaluación, la discusión se produce entre “antiguos pares” universitarios, investigadores o médicos, enfoque que también subyace en los proyectos aprobados por el Cime. En los primeros dos casos se formó un Consejo de Evaluación Universitaria y de Investigación que, en opinión de Thoenig (1996), obligó a la realización de discusiones argumentadas y a una evaluación con énfasis en los procesos, más que en los impactos.

Por otro lado, se percibía una tendencia a no querer perjudicar a nadie con los resultados de las evaluaciones, lo que se tradujo en información poco útil y una permanente tentación por autoevaluarse. Además, el parlamento francés no se había mostrado motivado por capacitarse en materia de evaluación para asumir una misión de control del Poder Ejecutivo. Sólo contaban con un Servicio de Evaluación de las Elecciones Científicas y Técnicas, con mínimas atribuciones y pocos recursos.

Nioche (1993) resumía los problemas enfrentados por la evaluación a partir del decreto de 1990 en los siguientes: concepción limitada de tipo gerencial, estructura administrativa compuesta por tres organismos controlados por el Ejecutivo y ausencia de una institución independiente especializada en evaluación, falta de participación del parlamento, procedimientos complejos en que los roles y las responsabilidades no estaban claramente delimitados, falta de desarrollos metodológicos propios para enfrentar el rechazo relativo de los extranjeros, trabajo disperso y de poco rigor, limitada utilización de sus resultados en el debate público y la toma de decisiones administrativa, entre otros. Consideraba que la institucionalización de la evaluación de la acción pública requería de una reforma del Estado, que modificara el funcionamiento del parlamento y los cuerpos de control. Trosa (1993) agregaba la existencia de un incremento de procedimientos y un déficit de aprendizajes, así como resaltaba la falta de conexión entre la evaluación y la administración interna.

Por su parte, Lascoumes y Setbon (1996) describían las evaluaciones realizadas como: cuidadosas en la selección plural de actores, con fuerte presencia de la administración y los cuerpos de inspección, una cultura

dominante de inspección y auditoría, un déficit generalizado de profesionalización, una amplia autonomía de las instancias de evaluación, y generalmente, la presencia de un actor administrativo que demandó la evaluación, forma parte de dicha instancia, y trata de jugar un rol central. Setbon (1998) agrega una débil implicación de los actores políticos que impide que las evaluaciones se traduzcan en un reposicionamiento del tema en la agenda y en su transformación, mientras que se manifiesta un amplio interés de los administradores que operan programas, un avance en materia de esclarecimiento de problemas y de aprendizaje social bajo el concepto de evaluación formativa.

En 1996 se crearon dos oficinas parlamentarias. La primera, referida a la evaluación de políticas públicas, para enfrentar la necesidad de evaluar, retrospectiva y prospectivamente, la acción gubernamental y también la del propio Parlamento; todo esto, nuevamente, en el marco de las directivas adoptadas por la Comunidad Europea. La segunda, dedicada a la evaluación de la legislación, de composición estrictamente política, con menos poder que la anterior, ya que si bien puede recurrir a expertos, carece de un ordenamiento administrativo y financiero que facilite su misión. Esta decisión permitió restablecer un cierto equilibrio en los contrapesos entre poderes del sistema francés, con cuatro riesgos importantes: una evaluación meramente política (no objetiva, que defienda intereses partidarios), sin suficiente garantía científica, que dispusiera de medios limitados, y apoyada por una voluntad política incierta. Ese año se publicó también una *Pequeña guía para la evaluación de políticas públicas*.

Hacia finales de los noventa, Francia vivió una cierta consolidación y madurez institucional de la evaluación, acompañada de un relanzamiento de la misma a nivel gubernamental y parlamentario. En noviembre de 1998, se produjo un nuevo decreto del Presidente de la República (núm. 98-1048), que creó el Consejo Nacional de Evaluación de Políticas Públicas (CNE) y manifiesta que “La evaluación de una política pública, en el sentido del presente decreto, tiene por objeto apreciar, en un cuadro interministerial, la eficacia de esta política, comparando sus resultados a los objetivos asignados y los medios empleados” (traducción libre). Se trata del regreso al concepto de evaluación *ex post* para el mejoramiento de la eficacia de las administraciones públicas, basado en la aplicación de procedimientos más rápidos de evaluación (máximo de 18 meses, salvo casos excepcionales). El decreto busca un equilibrio entre la participación plural de los interesados (*stakeholders*) al estilo americano y el control, bien francés, del proceso por la administración. Se insiste en la transparencia y publicidad de los resultados, así como en una

nueva apertura del CNE a los representantes parlamentarios y miembros del Consejo Económico y Social.

El CNE, compuesto de 14 personas de tiempo parcial, nombradas por el gobierno por periodos de tres a seis años, está encargado de definir el programa anual de evaluaciones a proponer al Primer Ministro, registrar los reportes de evaluación de los comités y ofrecer el servicio de consulta metodológica por solicitud. El comisariado del plan actúa como apoyo secretarial del CNE y publica los informes de evaluación, junto con los comentarios de este comité y de los ministros.

El procedimiento en marcha implica una toma de decisión respecto de lo que será evaluado (leyes, políticas, programas), seguida del nombramiento por el gobierno del presidente del dispositivo o comité de conducción de la evaluación, una instancia independiente compuesta en cada caso por 12 a 30 miembros designados por el Primer Ministro a título personal (profesionales, diputados locales, expertos, servidores públicos, representantes de la sociedad civil y grupos de interés, entre otros) con objeto de monitorear la evaluación (definir metas, realizar la planeación de la investigación, efectuar audiencias, visitas de trabajo, etcétera), interpretar e integrar el conocimiento validado y escribir el informe final. Para el cumplimiento de su objetivo, el dispositivo de evaluación es responsable de escuchar a los diversos actores implicados, previo a la discusión de los problemas específicos que serán analizados en el trabajo y de la coordinación del estudio preliminar que permitirá la elaboración de los términos de referencia. Posteriormente elige a los encargados de realizar la evaluación, que pueden ser órganos de la administración pública, consultorías privadas, equipos universitarios o de investigación.<sup>26</sup>

De esa forma puede decirse que en la evaluación participan cuatro tipos de actores (autoridades públicas, instancias de evaluación interna, comités de evaluación *ad hoc* y expertos profesionales), que cumplen tres roles bien dife-

<sup>26</sup> Como ejemplo de la composición del dispositivo de evaluación, el correspondiente a la evaluación de la Ley Quinquenal del Trabajo, el Empleo y la Formación Profesional contó con seis miembros de la Administración Pública (ministerio correspondiente y Ministerios de Educación Nacional, Asuntos Sociales y Presupuesto, delegado interministerial y comisariado general del Plan), y nueve personalidades calificadas (un experto, dos investigadores universitarios, dos diputados locales, un operador, un empresario y dos representantes de asociaciones). Esta instancia fue asistida por un redactor general y varios redactores auxiliares puestos a disposición por los ministerios involucrados (Cime-primer ministro-comisariado general del Plan, 1997).

Otro ejemplo relevante lo constituyó la evaluación de la política de lucha contra el Síndrome de Inmuno Deficiencia Adquirida (SIDA) 1994-2000. En este caso se buscó que la instancia no estuviera conformada por especialistas en la epidemia sino por investigadores o representantes de asociaciones. Se integró con participación de diferentes disciplinas de las ciencias sociales intervinientes en el campo: epidemiología, sociología, economía, derecho, medicina general, psiquiatría, medicina social y preventiva, además de los representantes de los administradores, miembros de salud, justicia y educación y el jefe de la dirección de SIDA. Su primer presidente renunció a los ocho meses por no poder garantizar la calidad científica del informe en el tiempo disponible (CNE-comisariado general del Plan, 2002).



renciados (una instancia que solicita la evaluación, otra que la coordina y un órgano científico o experto que la realiza), con mecanismos de coordinación entre ellos. De esta organización depende en buena parte el éxito del trabajo, mismo que está, a su vez, marcado por las características de quien preside la instancia de evaluación *ad hoc*.

En 1999 se consideraba que la evaluación se había vuelto una realidad en el país, aunque Francia seguía presentando un retraso en materia de rendición de cuentas de la gestión pública a la sociedad. Thoenig (1998) había caracterizado el dispositivo de evaluación francés como extremadamente centralizado y dependiente de los decisores gubernamentales, confrontado a dudas sobre su legitimidad y utilidad al no poder garantizar la eficiencia y la eficacia de los actos públicos; razón por la que se debía desmitificarla al mismo tiempo que difundirla.

Freyssinet (1998) señalaba que, a pesar de que la evaluación es estructurada y financiada por el Estado, la paradoja es que sus resultados son muy críticos, al poner el acento en las desigualdades, disfuncionamientos e ineficacias. Piensa que esto se explica por la tradición de independencia y deontología de quienes componen las instancias de evaluación, entre cuyos miembros aparecen fuertes contradicciones.

También en 1999 hay que resaltar la puesta en marcha por el parlamento de una Misión de Evaluación y Control del Gasto Público (MCE) y la creación de la Sociedad Francesa de Evaluación (SFE), con la integración de 250 universitarios, consultores y miembros de la administración pública, que ya llegaban a 300 en 2002.

Al cierre del siglo, nadie discutía la importancia de la evaluación, pero su arraigo era meramente institucional; por otro lado, la experimentación había sido totalmente desechada y sustituida por el “pluralismo a la francesa”, también el papel que originalmente cumplieron los economistas como evaluadores, quienes fueron reemplazados poco a poco por los sociólogos. El fuerte peso relativo del Ejecutivo sobre el Legislativo, así como su centralización histórica, había dificultado un desarrollo que tampoco había contado con una importante demanda ciudadana.

En agosto 2000 se produjo una circular del primer ministro sobre la evaluación de los contratos de Plan Estado-región que relanza los procedimientos impulsados por el comisariado general del Plan. El número de evaluaciones realizadas en Francia desde 1980 alcanzaba en 2002 la cifra aproximada de 1,500, con un amplio espectro de métodos y en la mayoría de los campos, incluyendo políticas locales (42 en total), ministeriales e interministeriales.

Hespel sintetiza el desarrollo metodológico e institucional que la evaluación ha logrado en el último decenio, pero advierte que las aportaciones logradas "...indican la necesidad de atemperar con pragmatismo algunas ambiciones iniciales" (Hespel, 2002: 51, traducción libre). Y es que se han difundido métodos y criterios de evaluación, se han formalizado los procesos en todos los ministerios, se ha desarrollado una *expertise* universitaria y administrativa, la Sociedad Francesa de la Evaluación ha realizado exitosos coloquios, se incrementó sustantivamente el número de evaluaciones interministeriales (15 en tres años), se ha fortalecido el papel del Parlamento en relación con las finanzas públicas a partir de la reforma de su ordenamiento orgánico en 2001, etcétera; pero, la cultura de gestión pública asociada a la evaluación está desigualmente desarrollada. Se trata de un déficit de prácticas (no de procedimientos formales), lo que no impide que tanto la izquierda como la derecha promuevan su fortalecimiento. Ha servido al menos para romper las barreras entre las instituciones que necesitan compartir la evaluación.

Como contraparte, los recursos que el CNE y el comisariado destinan a la evaluación siguen siendo muy limitados (unas 20 y 15 personas respectivamente en 2002, a las que se suman aproximadamente 150 miembros de las instancias de evaluaciones en curso, sin remuneración), la *expertise* se ha desarrollado pero alcanza apenas unas decenas o una centena de personas, los recursos financieros son insuficientes, y el CNE no cuenta con suficiente independencia del Ejecutivo, ya que depende del comisariado y necesita del acuerdo del primer ministro para solicitar una evaluación.

Thoenig (2002) percibe un desencanto de quienes esperaban un pilotaje más científico de los asuntos públicos, tanto a nivel gubernamental como del debate político. Algunos de los problemas que observa en el estancamiento de la evaluación se relacionan con su institucionalización formal exagerada (procedimientos), los obstáculos interpuestos para impedirla por los juegos de poder, la existencia de funcionarios y consultores que denominan evaluación a estudios que no lo son, su déficit de profesionalismo, la falta de recursos para su desarrollo y la reducida utilización de sus resultados. Se han multiplicado los excesos como la experimentación de los años sesenta en los EUA, la insistencia tatcheriana del *value for money* en Reino Unido y su institucionalización por decreto en Francia.

Sin embargo, esto no significa para el autor que, en determinadas condiciones, las evaluaciones no contribuyan a aportar "un poco de conocimiento razonado y riguroso" a los juicios que orientan la acción pública, tanto en la perspectiva de la gestión como en la dirección política del gobierno. En

particular, “el uso de la evaluación se aceleraría si existiera una comunidad profesional digna de ese nombre” y “uno o dos órganos instrumentalmente competentes instalados en el corazón de la maquinaria gubernamental y que gozaran de una especie de magistratura moral” (Thoenig, 2002: 458, traducción libre).

Thoenig identifica dos razones básicas para promover una mayor dosis de evaluación: nuevamente, el aprendizaje organizacional y de sus actores; y una mejor fundamentación de la autoridad política. La productividad y la eficiencia no entusiasman al ciudadano, tampoco éste valora de inmediato las ganancias en transparencia ni aprueba sus costos, lo que, en ausencia de sus referencias clásicas, ha restado fundamentación a la evaluación. Pero el Estado no sólo es cuestionado en términos de eficiencia, sino también en su capacidad para tratar los problemas públicos, lo que, para el autor, abre la segunda posibilidad para la evaluación: participar en la construcción de una nueva autoridad política basada en un gobierno compartido con la sociedad, facilitar la identificación de las partes que se deben involucrar, desarrollar mesas de debate sobre los problemas, etcétera.

Los cambios mencionados requieren transformaciones profundas en los servicios de *expertise* interna de los servicios centrales del Estado, cuya tendencia a la regulación tecnocrática y corporativista persiste desde la época de los cincuenta-sesenta en que se impuso la planificación y la racionalización de las elecciones presupuestales. Este enfoque es aún liderado por pequeños grupos elitistas (Ecole Polytechnique), cuyo monopolio del saber se les escapa cada vez más.

Thoenig subraya también que los contratos no garantizan el éxito y calidad de las evaluaciones, que los ministerios centrales enfatizan la medición de resultados más que de impactos y tienden a considerar que la información más sofisticada es mejor, que la toma de decisiones sigue siendo política por lo que la evaluación es sólo una fuente adicional de insumos que debe ayudar a los procesos de aprendizaje para pensar bien y ser oportunos. La paradoja es que, a pesar de estar tan ampliamente difundida, la evaluación sigue siendo poco usada como herramienta gubernamental.

## Comentarios finales

El capítulo nos muestra la promoción que los organismos internacionales hacen del tema de la evaluación, a pesar de la variabilidad de sus enfoques. Por su parte, los países desarrollados revisados nos aportan aprendizajes de gran utilidad:

EUA nos ofrece información sobre los costos y peligros de una excesiva experimentación. Reino Unido destaca la utilidad de un trabajo de preparación de materiales y capacitación de funcionarios a todos los niveles, antes y durante la institucionalización de procesos de evaluación.

Canadá, país en que ésta se desarrollaba como monopolio del Ejecutivo, nos advierte sobre el riesgo de transformar a la evaluación en un nuevo ritual que evite hacer críticas a los programas, no mida lo importante (la pertinencia de los programas y su impacto), no sugiera cambios sustantivos y no sea utilizada en los procesos decisorios. También sobre la importancia de realizar evaluaciones de políticas a nivel interministerial y de impulsar amplios procesos de participación social.

Francia nos enseña sobre los peligros del centralismo y la falta de involucramiento del Poder Legislativo, el exceso de preocupación por los métodos de evaluación en detrimento de los problemas de organización y funcionamiento del contexto organizacional, la necesidad de contar con organismos, evaluadores y un fondo de financiamiento autónomos y de realizar evaluaciones selectivas de programas, entre otros.

En el siguiente capítulo se analizará la situación que en materia de experiencias de evaluación de políticas y programas públicos prevalece en AL y México en particular.

## Capítulo 5

# Desarrollos empíricos de la evaluación en América Latina

### Introducción

COMO ya se mencionó, la mayoría de los países del mundo ha intentado aplicar la evaluación a partir de las dos últimas décadas del siglo pasado. AL, región en que persisten resistentes problemas de pobreza, desempleo, bajos salarios, salud, educación, vivienda, etcétera que sus gobiernos deben enfrentar con recursos muy limitados debido a los procesos de ajuste aplicados, cuenta con pocas experiencias de evaluación y de impacto aun limitado.

México presenta múltiples intentos de evaluación, más o menos dispersos a lo largo del siglo XX, y un importante esfuerzo de sistematización de los procesos, mediante la modalidad de evaluación externa, a partir del año 2001.

### América Latina: ejemplos escasos y parciales

En AL, durante los años cincuenta y sesenta en que se produjeron reformas sociales populistas (con Perón en Argentina, Batlle en Uruguay, etcétera), la evaluación se aplicó en forma mínima. En los setenta, su uso se incrementó en algunos programas de desarrollo, promovido por organismos multilaterales y fundaciones internacionales que financiaban programas de apoyo.

En épocas más recientes, múltiples trabajos señalan el limitado avance de la evaluación en los países de la región. Como ya se mencionó, Sulbrandt (1993) consideraba que la mayoría de las políticas y los programas públicos de AL no se evaluaban, especialmente los programas sociales, y que, salvo excepciones, la evaluación no había conducido a provocar ningún cambio en el manejo de los programas ni en el aprendizaje de los ciudadanos.

Por su parte, Wiesner (2000) afirma que la evaluación por resultados no ha logrado despegar debido a intereses particulares que conducen a la realización de evaluaciones espúreas, “precisas pero falsas” (pseudoevaluaciones) y débilmente vinculadas a los procesos de asignación de recursos, que impiden se convierta en un instrumento de reforma y modernización.

También Cohen (2001) manifiesta que no existe claridad sobre los beneficios provocados por la aplicación de programas para la reducción de la pobreza; lo que implica una crítica a la forma en que estos se han diseñado, pero también a la insuficiencia de las evaluaciones realizadas, especialmente en la medición de su impacto social.

Brown (2001) señala que en los países en desarrollo, AL y el Caribe en particular, los organismos ejecutores raras veces evalúan los resultados de su gestión, siendo ésta promovida casi siempre por los organismos financieros (BM, BID, etcétera) para satisfacer sus propios requisitos de responsabilidad fiscal. Plantea que la falta de evaluación de políticas y programas impide una mayor utilidad de los presupuestos por programas y por resultados y provoca importantes costos por falta de aprendizaje y repetición de errores.

De igual manera, Ospina (2001) considera que son pocos los casos en AL (cita como ejemplos a Chile, Colombia y Costa Rica) en los que se puede hablar de una relativa evaluación sistemática en la administración pública y de la utilización de sus resultados, en sus cuatro niveles: políticas públicas, programas públicos, organizaciones públicas y desempeño de los empleados públicos. Piensa que la evaluación de la gestión pública está dejando de ser un ejercicio técnico para transformarse en una herramienta política para resolver problemas de gobernabilidad y legitimidad, vinculada al enfoque de NPM. Así una mejor gestión del gasto público, afirma, mejora la imagen del Estado y aumenta su legitimidad, libera recursos que pueden ser aplicados a satisfacer necesidades de grupos sociales de menores ingresos, aumenta la equidad, etcétera. Sin embargo, no se reporta ningún caso de AL en que estos resultados deseables se presenten empíricamente. A pesar de la fortaleza de los esfuerzos chileno y colombiano, Ospina duda aún de que se logre promover la gestión orientada a resultados y el consecuente mejoramiento de la administración pública que constituye su objetivo.

Los factores que habrían desalentado a la evaluación son: el proceso presupuestal, el alto costo de las evaluaciones, la toma de decisiones centralizada, la falta de un interés continuado de parte de políticos y funcionarios

públicos, la ausencia de un método eficaz para determinar el éxito o fracaso de los programas, las dificultades para utilizar indicadores directos, la insuficiencia de equipos técnicos, el interés exclusivo en el control de insumos y la observación de las normas, el privilegio del volumen sobre la calidad y de la cobertura sobre los resultados y la satisfacción, entre otros (Brown, 2001; Cohen, 2001).

Para Ospina el desafío es generar una dinámica que motive a los líderes de una organización pública a demandar y usar información de manera explícita y sistemática para tomar decisiones gerenciales y políticas, y un diseño institucional que produzca los incentivos necesarios para que los usuarios vean la importancia de acceder a la información, ya que, según Repetto (2003), tampoco la ciudadanía muestra siempre el interés o las habilidades necesarias para utilizar plenamente la información generada en la evaluación.

En un reciente trabajo (Cunill y Ospina, 2003), las autoras consideran que su estudio es el primero, hecho en profundidad, en la región, ya que insisten en que en AL se le ha prestado muy poca atención a la evaluación de la gestión pública, en comparación con la existente en EUA y los países integrantes de la OCDE.

En su estudio analizan los casos de los países identificados por Ospina en el trabajo anterior y agregan el de Uruguay, en un intento por medir los actuales esfuerzos de evaluación de la gestión pública en los procesos de modernización y democratización del Estado y la administración pública en la región, en los cuatro casos en que existen algunas iniciativas sólidas y herramientas de evaluación en operación.<sup>27</sup> Identifican las diferencias y semejanzas de los diferentes sistemas diseñados y su impacto en el mejoramiento de dicha gestión, con la finalidad de difundir estas experiencias que pueden servir como oportunidad de aprendizaje para otros países. En el nivel meso, de evaluación de organizaciones y programas gubernamentales, destacan los avances del caso chileno.

De su trabajo extraemos los principales antecedentes de la evaluación de la gestión pública (que incluye a la evaluación de políticas y programas, pero también a la evaluación institucional y del desempeño individual) en los cuatro países citados (Colombia, Costa Rica, Chile y Uruguay), mismos que se resumen en el siguiente esquema:

<sup>27</sup> También reconocen avances menores en México, Brasil y Argentina, pero no los incluyen en el trabajo. Recientemente el CLAD ha agregado el estudio de nuevos casos nacionales (Argentina, Brasil y Honduras) que pueden consultarse en [www.clad.org.ve](http://www.clad.org.ve)

ESQUEMA I  
EVOLUCIÓN DE LA EVALUACIÓN EN AMÉRICA LATINA

<i>Fecha</i>	<i>País</i>	<i>Actividad</i>
1967-1968	Uruguay	Introduce el presupuesto por programas y la evaluación incisos del Presupuesto Nacional.
1974	Costa Rica	Encomienda al Ministerio de Planificación Nacional y Política Económica (Mideplan) la evaluación sistemática de los resultados de los planes, políticas y programas de desarrollo.
Década de los ochenta	Colombia	Se crea el Sistema de Evaluación y Seguimiento de Proyectos (Sisep), enfocado a los proyectos financiados por el BM y el BID, donde también participa el Banco de Proyectos de Inversión Nacional, encargado de las iniciativas del gobierno nacional y dirigido a la evaluación <i>ex ante</i> .
1987	Uruguay	Retoma su proceso al aprobar la Ley de Rendición de Cuentas, que incluye información económico-financiera sobre el cumplimiento de metas y objetivos.
1991	Colombia	La creación de un sistema de evaluación de la gestión y resultados de la administración pública, es encargada al Departamento Nacional de Planeación, como consecuencia de una reforma constitucional.
1992	Colombia	Se crea la División Especial de Evaluación y Gestión dentro del Departamento antes citado, institucionalizando las <i>evaluaciones ex post</i> .
1994	Colombia	Primera propuesta de Sistema Nacional de Evaluación de Resultados de la Gestión Pública (Sinergia), como herramienta para generar un cambio en la cultura de administración pública, abriendo la gestión a los ciudadanos, al ofrecerles amplia información ( <i>accountability</i> ). Este sistema careció del apoyo político necesario como consecuencia de la crisis de legitimidad que vivía el gobierno, pero fue validado por el BM y promovido como modelo ejemplar.
1994	Costa Rica	Su Plan Nacional de Desarrollo 1994-1998 crea un sistema de evaluación de la gestión y un área específica de Evaluación y Seguimiento, que se acompaña de un decreto del ejecutivo que da origen al Sistema Nacional de Evaluación (Sine). Este sistema surgió con un fuerte respaldo de la Presidencia de la República, mismo que ha mantenido a pesar de los cambios de administración.



<i>Fecha</i>	<i>País</i>	<i>Actividad</i>
1994	Chile	El gobierno inicia el impulso a la modernización de la gestión. El Banco Mundial (2004a) informa que se habría creado un sistema de Control de Gestión y Presupuestos por Resultados.
1995	Chile	Se impulsa la utilización voluntaria de indicadores de desempeño.
1995-2000	Uruguay	La evaluación recibió su mayor impulso a partir de la Ley de Presupuesto 1995-1999.
1997	Uruguay	Se crea el Sistema Informático de Evaluación de Gestión Pública (SEV).
1998	Uruguay	Se aprueba la Ley de Rendición de Cuentas.
1999	Uruguay	El SEV comienza sus operaciones; pero, a partir de estas fechas el sistema ha tendido a declinar.
1997-1999	Chile	Se produce el primer intento por articular sus instrumentos de evaluación (de metas ministeriales, desempeño de funcionarios, etcétera) en un sistema integrado, con la creación del Sistema de Evaluación de Programas Gubernamentales y la implantación de los Balances de Gestión Integral. Los indicadores de desempeño se vuelven obligatorios, pero su uso es interrumpido dos años después.
2000	Colombia	Sinergia logra el apoyo político suficiente para pasar a ser orgánicamente una Dirección de Evaluación de Gestión y Resultados.
2000	Costa Rica	La nueva Constitución incorpora la evaluación por resultados y la rendición de cuentas como principios inspiradores de la institucionalidad democrática.
2000	Chile	El uso de los indicadores de desempeño se vuelve a adoptar en forma obligatoria para todas las entidades públicas, después de ajustarlos metodológicamente. Conjuntamente se adoptan los mecanismos de evaluación restantes. Sin embargo, todos estos mecanismos no integran un sistema único y articulado, distribuyéndose su responsabilidad entre varias entidades del gobierno central (División de Control de Gestión, Secretaría General de la Presidencia, Ministerio de Planificación, etcétera). Su apoyo legal está basado en leyes, decretos y acuerdos entre los poderes Ejecutivo y Legislativo.

Fuente: Elaboración propia con base en Cunill y Ospina (2003).

La situación actual en estos países es la siguiente:

El sistema colombiano abarca los niveles macro (políticas) y meso (unidades ejecutoras) de la gestión pública, aunque fue concebido para el primero, en el que no ha logrado incidir para reorientar las políticas públicas. En su funcionamiento existen superposiciones de atribuciones, particularmente entre la Contraloría, el Ministerio de Hacienda y el equipo de Sinergia, quienes, en 2001, firmaron un convenio para resolver estos problemas. El sistema mantiene escasa vinculación con los poderes Legislativo y Judicial y se tienen planes para su descentralización a mediano plazo. Goza de un alto nivel de legitimidad porque se reconoce la necesidad de evaluar los resultados, aunque es poco conocido entre políticos, académicos y población en general.

En Costa Rica el diseño pretende evaluar la gestión pública en sus tres niveles (macro, meso y micro) para ofrecer información a la población que garantice la transparencia y legitimidad del sistema, así como el fortalecimiento democrático; pero hasta ahora sólo se ha evaluado parcialmente el nivel meso. Por otra parte, no existen mecanismos de coordinación formal entre los diversos organismos de evaluación y control.

Mientras tanto, en Chile, la evaluación se hace efectiva en los tres niveles de la gestión y cuenta con amplia legitimidad, pero tampoco existe una articulación formal entre los distintos mecanismos, los que se ha intentado complementar con la modernización de los servidores públicos, con un avance gradual de su descentralización y desconcentración a nivel regional y local, y la intención de incorporar a la ciudadanía al diseño y evaluación de las políticas y programas públicos. Recientemente se ha impulsado la adopción de un Sistema de Información para la Gestión Financiera del Estado (SIGFE).

Finalmente, Uruguay concentra su evaluación en el nivel meso (aunque algunas entidades quedan fuera del sistema), sin pretender evaluar las políticas públicas. El objetivo del Sistema de Evaluación de la Gestión Pública por Resultados (SEGPR) es proporcionar una herramienta para la toma de decisiones, que permita mejorar la asignación del gasto público y focalizar la acción en el ciudadano. Si bien existen cuatro organismos designados para participar en procesos de evaluación, no ha habido problemas de superposición debido a que, en 2001, dos de ellos aún no funcionaban y además sus atribuciones están claramente delimitadas por la ley. También se ha intentado, aunque sin resultados positivos, incorporar al sistema a las ONG que reciben subsidios públicos y cuentan con sus propios mecanismos de evaluación.

La comparación de los cuatro casos nacionales muestra que sus reformas han estado fuertemente influidas por el NPM; sin embargo, no son uniformes en diversos aspectos:

- Salvo el caso de Colombia que resolvió diseñar su sistema en forma independiente de toda experiencia previa, los demás países tomaron éstas en cuenta, particularmente la de Gran Bretaña en los casos de Costa Rica y Chile. El último, además, intercambió profesionales con Australia y Nueva Zelanda; en Uruguay también se sintió la influencia del NPM. Para Costa Rica fue relevante la experiencia de Colombia, y para Uruguay, la de Chile.
- Hay un grupo de países que tomó al proceso presupuestario como eje básico de la misma (modelo Presupuesto, en términos de Cunill y Ospina), modificando su énfasis en el control de los gastos por una mayor preocupación de la evaluación de los resultados medibles (Uruguay y Chile); mientras los dos restantes buscaron la creación de mecanismos que aumenten la capacidad institucional del Estado mediante la modernización de las organizaciones públicas, que refuercen sus decisiones políticas, más que las presupuestarias (modelo Plan, en la misma fuente mencionada).
- En Uruguay, Colombia y Costa Rica se crea un sistema nacional de evaluación, centralizado y dirigido desde un nivel estratégico del aparato estatal sobre el conjunto de la administración pública del país; en Chile se trata de un conjunto de continuas iniciativas relativamente independientes, sin la pretensión anterior.
- Los sistemas presentados fueron introducidos en forma gradual, mediante pruebas piloto, salvo en el caso uruguayo que se implantó simultáneamente, en un tiempo muy reducido, en todas las unidades Ejecutoras del Presupuesto Nacional, lo que dificultó la acumulación de aprendizajes.
- Su incorporación fue acompañada en todos los casos de procesos de capacitación, realización de seminarios, talleres, difusión de otras experiencias, distribución de manuales, instructivos, guías metodológicas, libros, videos, etcétera.
- Costa Rica y Uruguay recibieron apoyo técnico del BID, Colombia del BM y Chile de ambos.
- En los países analizados, el sistema cubre sólo el nivel central de gobierno, salvo Chile, que muestra algunos avances adicionales a nivel regional y Costa Rica que incluye algunas instituciones descentraliza-

das. Se evalúan organizaciones y programas seleccionados, mediante consultoría externa y algunos monitoreos a cargo de los ministerios respectivos.

- Mientras Uruguay ha definido un gran número de indicadores que se consideran poco efectivos, Chile y Colombia han aprendido a limitar su número y mejorar su diseño, aumentar la calidad de la información y facilitar su utilización.
- La legitimidad en el diseño de los anteriores depende de la participación ciudadana en esta etapa, asunto que Costa Rica comienza a considerar. Su debate requeriría de una evaluación de tipo pluralista como la propuesta por Monnier.
- El mecanismo empleado para informar resultados es principalmente la presentación de informes al Congreso. En el caso colombiano se realiza anualmente e incluye un resumen de todas las evaluaciones realizadas, el análisis del grado de cumplimiento del plan y propuestas de ajustes. En Chile, los presenta cada institución, utilizando como instrumento el Balance de Gestión Integral, que sintetiza los resultados de las evaluaciones. Uruguay presenta similar información en el Documento de Rendición de Cuentas. Finalmente, Colombia presenta, cada cuatro meses, al Consejo Nacional de Política Económica y Social, información del seguimiento realizado, y anualmente, de la evaluación efectuada.
- Como incentivo a las buenas prácticas, los cuatro países utilizan la difusión pública de los resultados de las evaluaciones de desempeño institucional. Las vías de divulgación generalizada son las páginas web institucionales; Costa Rica utiliza también conferencias de prensa sobre el tema. En Uruguay, la principal forma de acceso a esta información es la prensa, pero ésta le ha prestado muy poca atención al tema. Esta limitación se agudiza debido al lenguaje excesivamente técnico en que se redactan los documentos de difusión. En el esquema chileno se vinculan los logros alcanzados con incentivos de diversos tipos, como los Programas de Mejoramiento de Gestión (PMG), que relacionan el nivel de desempeño con las decisiones sobre incrementos salariales y los resultados de las evaluaciones de programas gubernamentales, que se utilizan para definir su marco presupuestario, a través de un sistema de “premios” y “castigos”. Su utilización puede convertirla en una actividad amenazante para la gerencia, que genera respuesta defensivas, en lugar de coadyuvar a mejorar el rendimiento, por

lo que parece más efectivo el uso de incentivos simbólicos de los otros países.

- Si no se hace uso de la información con algún tipo de consecuencias directas, como ha sucedido en los casos de Colombia y Uruguay cuando el resultado no es el esperado, y no se realiza un aprendizaje ciudadano de los instrumentos de evaluación, el sistema no se toma con seriedad. En los dos países citados y Chile, aunque la información se publique en Internet, no hay evidencia de que los ciudadanos accedan a ella, lo que Costa Rica ha resuelto, en parte, con sus publicaciones en la prensa.
- En los últimos cinco años se han fortalecido las organizaciones de la sociedad civil especializadas en la evaluación de la gestión pública, lo que no sucede en el caso de los legisladores. Ellos parecen ser actores privilegiados sólo bajo el modelo Presupuesto, sin embargo, enfrentan fuertes limitaciones tanto en Chile como en Uruguay para afectar la asignación de recursos; mientras que en el modelo Plan ese lugar lo ocupa la Presidencia.
- En cuanto a los logros reales de los sistemas de evaluación descriptos, aún queda mucho por hacer para que se acerquen a lo esperado, sin embargo, recién están en proceso de implementación:
  - Colombia ha avanzado en la generación de una cultura orientada a resultados y la rendición de cuentas y logró que los valores en que se sustenta la evaluación cambiaran de la eficiencia del gasto a los resultados e impactos. También presenta un desarrollo importante de los instrumentos, una rápida consolidación del sistema y una definición más clara de responsabilidades; pero, las conclusiones de las evaluaciones rara vez son incorporadas en los procesos decisorios.
  - En Costa Rica el sistema ha contribuido a promover la conducción estratégica del gobierno y la rendición de cuentas, pero no a desarrollar capacidades en materia de programación presupuestaria, otro de sus objetivos.
  - El incremento de la eficiencia de la gestión pública, la medición de sus resultados y el aumento de la transparencia, constituyen los principales logros chilenos, a pesar de su falta de integración. Durante el año 2002, se habrían evaluado 12 de sus programas, ocho de los cuales persiguen objetivos sociales.
  - En el caso uruguayo, está funcionando un sistema de indicadores de desempeño, pero no se ha alcanzado el nivel de una verdadera gestión

del desempeño, donde la evaluación se transforme en aprendizaje y se utilice en la toma de decisiones, principalmente a nivel estratégico. No existe ningún mecanismo destinado a incentivar su uso. Tampoco su Parlamento cuenta con información suficiente para analizar el cumplimiento de metas.

Contrario a lo esperado por las autoras, el sistema nacional y coherente de Uruguay parece haber sido menos efectivo que las diversas iniciativas inconexas de Chile. En realidad, reconocen que no cuentan aún con evidencia suficiente en ninguno de los países que instalaron sistemas de evaluación para asegurar que hayan sido afectadas positivamente la calidad de los servicios o la productividad del sector público.

Entre los desafíos comunes a estos países figuran el mejoramiento de los sistemas de información, así como su calidad y confiabilidad, las estrategias de difusión de resultados y el desarrollo o consolidación de incentivos vinculados al uso de la información generada sobre los resultados alcanzados.

A pesar de ello, Cunill y Ospina concluyen que existen motivos para ser optimistas, ya que hay suficiente consenso entre los funcionarios y gerentes públicos sobre la importancia de orientar su gestión a resultados, lo que constituye el primer paso hacia una cultura de nueva gestión pública y se percibe una notoria diferencia respecto de la situación prevaleciente hace cinco años.

Complementariamente, el BM (2004a) señala que en Chile, al igual que en los EUA, se ha implementado un esquema gubernamental integral de M&E con base en la Administración Basada en Resultados y se han capacitado 180 profesionales de diferentes instituciones y ministerios en materia de evaluación de programas. A mediados de 2004, dicho país puso también en marcha, con asistencia del BID, un sistema que incluye la selección de los gerentes públicos por un Consejo de Alta Dirección Pública, cuyos integrantes son nombrados por el Senado, y que, en materia de evaluación, prevé su realización por una entidad independiente de los poderes Ejecutivo y Legislativo, similar al Rekenkamer en Holanda. Weissbluth (2004) consideró que se trata de la reforma más profunda realizada en AL en las últimas décadas.

Agrega que Bolivia intentó un sistema de evaluación de desempeño, Sistema de Seguimiento y Evaluación de la Gestión Pública por Resultados (Siser), pero que éste tuvo que dejar de funcionar. En cambio, en Brasil, la mayor parte de los programas de su Plan Plurianual se evalúan anual-

mente y su informe es entregado al Congreso para considerarlo en el ciclo presupuestal; sin embargo, habría pocas evidencias de que sean tomados en cuenta.

### México: una sistematización largamente postergada

Desde el siglo pasado muchos programas públicos, y especialmente sociales, han sido desarrollados por diversas secretarías de estado (Salud, Educación Pública, Desarrollo Social, etcétera), organismos descentralizados (Instituto Mexicano del Seguro Social, Instituto de Seguridad y Servicios Sociales de los Trabajadores del Estado, Instituto Nacional Indigenista, etcétera) y empresas de participación estatal mayoritaria (Comisión Nacional de Subsistencias Populares, Conasupo; Leche Industrializada Conasupo, Liconsa; Distribuidora Conasupo, Diconsa, etcétera), en México.

A partir de 1917 se realizaron intentos por crear un organismo que midiera, siquiera parcialmente, los resultados de la gestión gubernamental, poniendo énfasis en el manejo de los recursos. Sin embargo, la lógica sexenal hizo que cada administración llegara con su nuevo personal público y reorientara los programas gubernamentales para ponerles su sello propio, provocando el desinterés por evaluar lo hecho por la anterior, situación que prevaleció hasta iniciado el siglo XXI.

De esta forma, lo realizado en materia de evaluación fue poco, parcial y estuvo a cargo de diversas instancias como la Secretaría del Patrimonio Nacional, la Secretaría de la Presidencia o la de Hacienda y Crédito Público; desconociéndose realmente los niveles de eficiencia y de impacto de los programas gubernamentales en los problemas que pretendían resolver.

Los intentos de reformas, especialmente dirigidos a los procesos de planeación de los que la evaluación formó parte, comienzan en la administración de López Portillo, y su origen podría ubicarse en los trabajos realizados, desde 1965, por la Comisión de Administración Pública (CAP), a la que pertenecieron tanto él como varios colaboradores suyos.

En el siguiente esquema se resumen los principales antecedentes vinculados al tema de la evaluación de la gestión gubernamental en general y de los programas de desarrollo social en particular:<sup>28</sup>

<sup>28</sup> Por razones de espacio no se explicitan avances específicos en los sectores de salud y educación.

## ESQUEMA 2

## ANTECEDENTES DE LA EVALUACIÓN EN MÉXICO

<i>Fecha</i>	<i>Actividad</i>
Hacia finales de los sesenta	<ul style="list-style-type: none"> <li>• <i>Evaluaciones independientes:</i></li> </ul> <p>Las primeras evaluaciones fueron desarrolladas en forma independiente por universidades y centros de investigación y continuaron realizándose hasta la última década del siglo XX, en pocas ocasiones y respondiendo a iniciativas de los propios investigadores. Así encontramos el primer antecedente, en que la Universidad Nacional Autónoma de México (UNAM) realizó los esfuerzos dirigidos a la implantación de un Presupuesto por programas.</p>
1974, 1977	<ul style="list-style-type: none"> <li>• <i>Congreso de la Unión:</i></li> </ul> <p>Debido a las modificaciones introducidas en los años señalados, el Congreso de la Unión adquirió algunas facultades respecto a los organismos paraestatales, además del control general de sus ingresos y egresos, pudiendo citar a los directores y administradores de los mismos para que informaran cuando se discutiera una ley o se estudiara un negocio concerniente a su ramo y también para que, cuando lo pidiera una cuarta parte de los miembros de la Cámara de Diputados y la mitad de la de Senadores, se integraran comisiones para investigar el funcionamiento de dichas entidades, y para poner en conocimiento del Ejecutivo el resultado de sus investigaciones.</p>
1976-1982	<ul style="list-style-type: none"> <li>• <i>Administración pública:</i></li> </ul> <p>El trabajo efectuado en la UNAM fue continuado, durante el sexenio del licenciado López Portillo, quien creó la Secretaría de Programación y Presupuesto (SPP), encargada de la planeación, programación, presupuestación, control y evaluación del gasto público. La SPP contó con una Subsecretaría de Evaluación de los resultados de los programas gubernamentales, integrada sólo por una Dirección General, que sin mucha claridad por la falta de antecedentes, pretendió ir más allá de la vigilancia del gasto, identificando y corrigiendo errores, replanteando estrategias, etcétera, para lo que debía diseñar e implantar un sistema nacional de evaluación; sin embargo, sus funciones se duplicaron con las de la Coordinación de Evaluación, dependiente de la Secretaría de Presidencia, creada en 1977, y sus esfuerzos se diluyeron. La Subsecretaría de Evaluación acabó realizando estudios parciales y de poco impacto en el proceso.</p>
1978	<ul style="list-style-type: none"> <li>• <i>Congreso de la Unión:</i></li> </ul> <p>Se fortaleció al Poder Legislativo en su control de las acciones del Ejecutivo, al aprobarse la Ley Orgánica de la Contaduría Mayor de Hacienda (CMH), órgano cuyos antecedentes pueden rastrearse hasta la administración colonial del siglo XV. La CMH fue el órgano técnico de la Cámara de Diputados (estuvo vigente hasta 1999), para el cumplimiento de sus facultades exclusivas como revisor del manejo de fondos de la Cuenta Pública del gobierno federal.</p> <p>Debía rendir un informe previo a la Comisión de Presupuesto y Cuenta Pública de la Cámara dentro de los 10 primeros días de noviembre,</p>



<i>Fecha</i>	<i>Actividad</i>
	siguientes a la presentación de la Cuenta Pública y un informe definitivo a la misma Cámara sobre el resultado de la revisión, en los 10 primeros días de junio del año siguiente a su recepción (art. 74 constitucional). En caso de presentarse problemas se determinaban y fincaban las responsabilidades correspondientes.
1980	<ul style="list-style-type: none"> <li>• <i>Participación ciudadana:</i> El gabinete instituye el mecanismo de participación y consulta a la ciudadanía, que tíbiamente había empezado a impulsar en la llamada “reforma de ventanillas”, cuyo único logro fue la simplificación de algunos trámites. Sin embargo, no se avanzó más allá de la definición de sus objetivos.</li> </ul>
1980	<ul style="list-style-type: none"> <li>• <i>Administración pública:</i> Se publica el primer Plan Global de Desarrollo y Miguel de la Madrid, al frente de la SPP, transforma las funciones de la Subsecretaría de Evaluación, responsabilizándola de preparar la información para la elaboración de la Cuenta Pública.</li> </ul>
1982	<ul style="list-style-type: none"> <li>• <i>Administración pública:</i> El énfasis en la planeación del sexenio anterior no se abandona pero, la evaluación adquiere una centralidad que no había gozado nunca, para responder a la renovación moral ciudadana, cuando el presidente Miguel de la Madrid (1982-1988) crea la Secretaría General de la Federación (Secogef), con carácter de globalizadora y atribuciones normativas para el resto de los organismos gubernamentales, encargada de ejercer la fiscalización, el control interno global y la evaluación sobre la administración pública, a través de las contralorías internas de las dependencias y entidades. Dicha secretaría debía cumplir sus funciones al interior de la administración pública sin entrar en conflicto con la CMH, que lo hacía desde otro poder, el Legislativo. También para evitar estos, se disolvió la Coordinación de Evaluación en la Presidencia.</li> </ul>
1983	<ul style="list-style-type: none"> <li>• <i>Administración pública:</i> Frente a la severa crisis económica prevaleciente, la SPP pierde su énfasis en la planeación para pasar a encargarse de los programas de estabilización, saneamiento de las finanzas públicas, contención del gasto, etcétera. En este nuevo encuadre las funciones de evaluación se trasladaron completamente a la Secogef; sólo permaneció en la SPP la contabilidad gubernamental y el control sobre las adquisiciones, para cumplir con la elaboración de la Cuenta Pública. En la propia Secogef también se vivió el desvío de funciones de sus comisarios y delegados, que acabaron trabajando en la liquidación o fusión de empresas o en los recortes de personal, muy lejos de combatir la corrupción.</li> </ul>
1984	<ul style="list-style-type: none"> <li>• <i>Administración pública y consulta ciudadana:</i> Mediante la modificación del artículo 26 constitucional, fue puesto en marcha el Sistema Nacional de Planeación Democrática, a partir de las orientaciones de la</li> </ul>

## ESQUEMA 2 (Continuación)

Fecha	Actividad
	Ley de Planeación y el Plan Nacional de Desarrollo, de 1983. Trataba de incorporar el carácter público y consensual del mismo, así como la participación de las entidades federativas en la planeación del desarrollo. Indicaba el mecanismo de consulta popular para recoger la participación social en las etapas de formulación, control y evaluación, el régimen de obligatoriedad de la planeación dentro del Sector Público Federal, el de concertación con otros niveles de gobierno y el inductivo para el ámbito no gubernamental.
1985	<ul style="list-style-type: none"> <li>• <i>Administración pública:</i> Se realizó la primera y única reunión de evaluación del programa de simplificación administrativa de trámites, mismo que buscaba elevar la productividad y eficiencia del aparato estatal. Se trató de una mera descripción de avances.</li> </ul>
1985	<ul style="list-style-type: none"> <li>• <i>Auditorías externas:</i> Se amplía la utilización del mecanismo de contratación de despachos privados, que ya se había empleado en el sexenio previo, para la realización de auditorías sobre aspectos específicos de la gestión pública.</li> </ul>
1986	<ul style="list-style-type: none"> <li>• <i>Asesoría sobre evaluación:</i> Diconsa solicitó al Departamento de Administración Pública del Centro de Investigación y Docencia Económicas (CIDE) A.C. una propuesta para evaluar su impacto socioeconómico y político. La misma fue presentada pero se desconoce si llegó a aplicarse.</li> </ul>
1988	<ul style="list-style-type: none"> <li>• <i>Administración pública:</i> Al finalizar el sexenio, la Secogef había dedicado mucho tiempo a la coordinación; su trabajo se había centrado más en la sanción de irregularidades que en su prevención; muchas áreas de evaluación habían desaparecido con la crisis económica; y, sólo en los últimos meses, se realizaron reuniones de autoevaluación de las entidades y dependencias gubernamentales, en que éstas se limitaron a presentar sus acciones en forma positiva y sin la menor autocrítica.</li> </ul>
1988	<ul style="list-style-type: none"> <li>• <i>Administración pública:</i> El nuevo presidente Carlos Salinas de Gortari (1988-1994) pone en marcha el Programa Nacional de Solidaridad (Pronasol), manejado desde la SPP, a pesar de que ya no se trataba de una estrategia administrativa, sino de la forma en que se busca resolver el mayor problema social: la pobreza.</li> </ul>
1989	<ul style="list-style-type: none"> <li>• <i>Administración pública:</i> Se creó el Programa General de Simplificación de la Administración Pública Federal, que recogía algunos planteamientos del NPM e incluía entre sus objetivos: consolidar una cultura administrativa que propiciara la eficiencia y la productividad en la gestión de las instituciones gubernamentales, promover la evaluación de</li> </ul>

Fecha	Actividad
	la calidad y la transparencia en la prestación de los servicios públicos y prevenir y combatir la corrupción. La responsable de aplicarlo era la entonces Secogef.
1990	<ul style="list-style-type: none"> <li>• <i>Administración pública:</i> Se crea la Comisión Nacional de Derechos Humanos (CNDH), como órgano desconcentrado de la Secretaría de Gobernación, tomando como modelo el <i>ombudsman</i> escandinavo o el “defensor del pueblo” español, “...como instrumento de la sociedad para vigilar y auxiliar a las autoridades en la observancia de la ley y en la protección de las garantías individuales” (Aguilar, 2000d: 170-171).</li> </ul>
1990	<ul style="list-style-type: none"> <li>• <i>Tesis sobre medición de la pobreza:</i> El doctor Miguel Székely Pardo (actual subsecretario de Planeación, Prospectiva y Evaluación de la Sedesol), presenta su tesis “Identificación de la pobreza global y rural: un primer paso para incrementar el grado de acierto de los programas públicos” para obtener el grado de maestro en políticas públicas por el Instituto Tecnológico Autónomo de México (ITAM). El mismo contiene como estudio de caso un acercamiento hacia la evaluación de los programas a cargo de Liconsa. Concluye en la necesidad de focalizar los apoyos hacia la población rural pobre.</li> </ul>
1991	<ul style="list-style-type: none"> <li>• <i>Contraloría social:</i> En ese mismo sexenio afloran los primeros atisbos de una contraloría social en México, mismos que inician al interior del Pronasol, asumiendo el principio de corresponsabilidad, como vocales de control y vigilancia, elegidos por los miembros de los comités de solidaridad, a su vez, elegidos por las comunidades beneficiarias del programa. Se trata de órganos transitorios, limitados al control <i>ex post</i> de los programas sociales, sin respaldo legal, personalidad jurídica ni especificación taxativa de atribuciones, por lo que obviamente no podían aplicar sanciones en forma directa, sino sólo informar a la Secogef de las irregularidades que identificaran. También había la confusión de que fungían como delegados del Sistema Integral de Control y Evaluación de la Gestión Pública y del Sistema Nacional de Quejas y Atención a la Ciudadanía, a cargo de la Secogef (Cunill, 2000).</li> </ul>
1992	<ul style="list-style-type: none"> <li>• <i>Administración pública:</i> Se disuelve SPP y sus funciones son transferidas a la SHCP; también se crea la Secretaría de Desarrollo Social (Sedesol). Después de modificarse el artículo 102 de la Constitución, se otorga el carácter de organismo descentralizado con personalidad jurídica y patrimonio propio a la CNDH. Este último constituye uno de los organismos que ayudan al equilibrio de poderes y sus contrapesos, creado para la evaluación de un aspecto concreto.</li> </ul>
1993	<ul style="list-style-type: none"> <li>• <i>Evaluación externa:</i> La Organización Internacional del Trabajo (OIT) encarga al doctor Miguel Székely Pardo la realización de una evaluación al Pronasol. El autor publica un documento de trabajo “El Programa Nacional de Solidaridad en México” que incluye un capítulo titulado “Hacia una evaluación del Programa Nacional de</li> </ul>

## ESQUEMA 2 (Continuación)

Fecha	Actividad
	Solidaridad", en que reconoce la inexistencia de elementos en la literatura disponible para responder a las preguntas que exceden a la información sobre el gasto ejercido y realiza un ensayo sobre los diferentes componentes del programa, con énfasis en la descripción de experiencias con el componente productivo en el estado de Oaxaca.
s/f	<ul style="list-style-type: none"> <li>• <i>Evaluación externa:</i> El Colegio de México propone una primera metodología general para la evaluación de Pronasol, coordinada por José Luis Reyna y Alberto García Rocha.</li> </ul>
1994	<ul style="list-style-type: none"> <li>• <i>Administración pública:</i> La Secogef adquiere responsabilidades en materia de desarrollo administrativo y se convierte en Secretaría de la Contraloría y el Desarrollo Administrativo (Secodam).</li> </ul>
1995	<ul style="list-style-type: none"> <li>• <i>Administración pública:</i> El gobierno del licenciado Zedillo publicó el Programa para la Modernización de la Administración Pública Federal 1995-2000 (Promap), con la intención de desarrollar una nueva cultura del servicio público, guiada por criterios de eficacia, eficiencia y combate a la corrupción, y fundamentada en el NPM. Dicho programa identificaba entre los principales problemas a resolver: las deficiencias en los mecanismos de medición y evaluación del desempeño de la gestión pública, para los que se proponía el diseño de un sistema basado en objetivos claros y metas cuantificables, indicadores de desempeño, la revisión y adaptación del Sistema Integral de Información por parte de las Secretaría de Hacienda y Crédito Público (SHCP) y Secodam y el fortalecimiento y mayor independencia de los órganos internos de control preventivo y correctivo.</li> </ul>
1995-2000	<ul style="list-style-type: none"> <li>• <i>Administración pública:</i> La SHCP fue encargada de estudiar la posibilidad de reformar el sistema presupuestal, enfocándolo a la evaluación de resultados de la gestión pública, a través de la aplicación de un sistema de indicadores. Complementariamente, la Secodam se encargaría de los aspectos vinculados al control de la calidad, la atención al cliente y la lucha contra la corrupción. De esta forma, al igual que en los casos de Uruguay y Chile, el presupuesto se volvía el instrumento de la reforma del Estado.</li> </ul>
1995-2000	<ul style="list-style-type: none"> <li>• <i>Administración pública:</i> La SHCP construyó la denominada Reforma al Sistema Presupuestario, que proponía la integración de una Nueva Estructura Programática (NEP) y de un Sistema de Evaluación del Desempeño (SED) basado en resultados, con el fin de otorgar un amplio margen de libertad a los gerentes públicos. Pero, la propuesta anterior se instrumentó muy parcialmente, logrando sólo incorporar el discurso</li> </ul>

<i>Fecha</i>	<i>Actividad</i>
1995-2000	<p>gerencialista sobre la reforma de la administración pública e introducir un poco más de claridad y transparencia en las cuestiones presupuestales (Arellano, 2003).</p> <ul style="list-style-type: none"> <li>• <i>Administración pública:</i> Las aportaciones del Promap, en otros ámbitos, incluyeron la creación del Sistema Nacional de Quejas, Denuncias y Atención a la Ciudadanía, que promueve la participación de la última; y la reforma de un conjunto de leyes que permitieron modificar el funcionamiento de los órganos de Control Interno de la Administración Pública Federal. Efectivamente, estos órganos dependían directamente de los titulares de cada institución, formando parte de las oficialías mayores o áreas similares, y se encargaban de realizar funciones de control y vigilancia de tipo correctivo, así como evaluaciones enfocadas a la gestión de los procesos. A partir de las modificaciones introducidas, los mismos pasaron a ser designados y removidos por la Secodam, de la cual dependen funcional y jerárquicamente, logrando autonomía respecto de las entidades en las que desarrollan su labor, y redireccionando sus evaluaciones hacia los resultados de la gestión.</li> </ul>
1995-2000	<ul style="list-style-type: none"> <li>• <i>Evaluaciones institucionales:</i> Hacia la segunda mitad de los noventa se realizaron las primeras evaluaciones institucionales, de tipo interno, a programas de Liconsa y Diconsa.</li> </ul>
1997	<ul style="list-style-type: none"> <li>• <i>Evaluaciones externas:</i> Se inició el proyecto de evaluación externa más ambicioso solicitado por el gobierno mexicano, el del Programa de Educación, Salud y Alimentación (Progres, hoy Oportunidades).</li> </ul>
1999	<ul style="list-style-type: none"> <li>• <i>Congreso de la Unión:</i> La CMH, luego de la modificación de los artículos 73, 74, 78 y 79 de la Constitución, fue sustituida en sus funciones por el órgano técnico, Auditoría Superior de la Federación (ASF), de la Cámara de Diputados del Congreso de la Unión, que continúa vigente a la fecha. Dicha modificación se realizó con la intención manifiesta de crear una instancia superior de revisión y auditoría, de carácter imparcial y que permitiera superar las limitaciones de su antecesora. Entre las últimas se planteaban las limitaciones impuestas a sus facultades para la revisión de cuentas públicas, la extemporaneidad de la emisión de resultados de sus revisiones, la ausencia de una garantía legal plena de su autonomía de gestión frente al Ejecutivo, la confusión generada por su propia denominación, el poco conocimiento por la ciudadanía de sus atribuciones y trabajos, etcétera.</li> </ul>
2000	<ul style="list-style-type: none"> <li>• <i>Evaluaciones externas:</i> Se difundieron los Términos de Referencia para la Evaluación Externa del Fondo de Coconversión Social 1999-2000, a cargo del entonces Instituto Nacional de Solidaridad (Insol).</li> </ul>

## ESQUEMA 2 (Continuación)

Fecha	Actividad
Diciembre de 2000	<ul style="list-style-type: none"> <li>• <i>Congreso de la Unión:</i> Se aprueba la Ley de Fiscalización Superior de la Federación que respalda jurídicamente a la ASF, la que cuenta con autonomía técnica, de gestión, material y financiera, para decidir respecto de su organización interna, funcionamiento y resoluciones. A diferencia de la CMH, su titular y funcionarios tienen la obligación de guardar secreto de su trabajo hasta tanto rindan los informes correspondientes (<a href="http://www.asf.gob.mx">www.asf.gob.mx</a>).</li> </ul>
2000	<ul style="list-style-type: none"> <li>• <i>Administración pública:</i> La Sedesol creó una subsecretaría encargada de la evaluación de los programas que se encuentran bajo su responsabilidad.</li> </ul>
2000	<ul style="list-style-type: none"> <li>• <i>Evaluación externas solicitadas por el Congreso de la Unión:</i> El Congreso comenzó a exigir a partir de 2001 la evaluación externa anual de los programas financiados con recursos fiscales, sujetos a reglas de operación.</li> </ul>
Inicios del 2001	<ul style="list-style-type: none"> <li>• <i>Administración pública:</i> Se instaló la Comisión Intersecretarial para la Transparencia y el Combate a la Corrupción, que además de su obvio objetivo, busca elevar los niveles de eficiencia, eficacia y calidad en el servicio público. En el mes de abril de ese mismo año se aprobó el Programa Nacional para la Transparencia y el Combate a la Corrupción, que entre sus líneas estratégicas incluye: mecanismos de rendición de cuentas, de supervisión y control, y de participación ciudadana.</li> </ul>
Junio de 2001	<ul style="list-style-type: none"> <li>• <i>Administración pública:</i> Se implementó el Programa de Corto Plazo para el Fortalecimiento del Control Interno dentro de las dependencias y entidades de la administración pública federal, con el objetivo de promover la autonomía de gestión de las instituciones y reducir la discrecionalidad en el manejo de los recursos, así como permitir una mayor transparencia en el ejercicio del presupuesto.</li> </ul>
2002	<ul style="list-style-type: none"> <li>• <i>Información para la evaluación por la sociedad civil:</i> Se aprueba en el mes de junio la Ley Federal de Transparencia y Acceso a la Información Pública Gubernamental y la Comisión Federal de Mejora Regulatoria (Cofemer) se convierte en el primer órgano público federal que rinde cuentas a la ciudadanía sobre su gestión.</li> </ul>
2003	<ul style="list-style-type: none"> <li>• <i>Administración pública:</i> La Secodam se transforma en Secretaría de la Función Pública (SFP), continuando con su responsabilidad de fortalecer los sistemas de gestión y control, combatir la corrupción y fomentar la transparencia de la administración pública Federal, entre otros (<a href="http://www.sfp.gob.mx">www.sfp.gob.mx</a>).</li> </ul>

<i>Fecha</i>	<i>Actividad</i>
2003	<ul style="list-style-type: none"> <li>• <i>Administración pública:</i> Se aprueba la Ley General de Desarrollo Social (LGDS) que prevé la creación de un Consejo Nacional de Evaluación integrado por ocho personas, seis de las cuales deben ser investigadores académicos que cumplan determinados requisitos que garanticen su calidad, y reconoce a la Contraloría Social como el mecanismo en que los beneficiarios verifican la aplicación de recursos y el cumplimiento de metas.</li> </ul>
2004	<ul style="list-style-type: none"> <li>• <i>Organizaciones de la sociedad civil:</i> Se publica la Ley Federal de Fomento a la Actividades Realizadas por Organizaciones de la Sociedad Civil, que en su art.6, inc. XI, promueve su mayor participación en los procesos de políticas públicas, en particular en relación al seguimiento de las acciones públicas realizadas.</li> </ul>

Fuente: Elaboración propia.

De esta forma, durante la mayor parte del siglo pasado, a nivel gubernamental, más que evaluaciones del impacto de las políticas y programas implementados, encontramos mecanismos de control jurídico y financiero del gasto (actividades de contraloría) y seguimientos de avance de metas. Cuando existían recortes presupuestales, las primeras actividades que se reducían eran las de evaluación, consideradas menos útiles que las sustantivas.

En resumen, puede decirse que, hasta el año 1999, las actividades de control y evaluación de la administración pública eran ejercidas por la SHCP, la entonces Secodam y la CMH de la Cámara de Diputados. Se trataba fundamentalmente del control del gasto público, la legalidad de los procesos, el combate a la corrupción, etcétera, pero no se tiene conocimiento de que estas instancias evaluaran algún programa gubernamental, menos social, en términos de sus resultados o impactos.

La SHCP era, y continúa siendo, el órgano financiero del Estado federal y se encarga de los asuntos relativos a los ingresos de la Federación y entidades paraestatales; proyecto y cálculo de egresos del gobierno federal y de la administración pública paraestatal; planeación, presupuestación, contabilidad y evaluación; entre otras.

La Secodam era la encargada de planear, organizar y coordinar el sistema de control y evaluación gubernamental, la inspección del gasto público federal en congruencia con su presupuesto de egresos, el establecimiento de bases para regular las auditorías a dependencias y entidades de la administración pública y efectuarlas, el control de que el enriquecimiento de los funcionarios

estuviera debidamente justificado, que las obras del gobierno avanzaran al ritmo acordado y tuvieran la calidad convenida, etcétera. Como ya se mencionó, en el sexenio del licenciado Zedillo se le agregó la responsabilidad del desarrollo administrativo de la administración pública.

La CMH ejercía el control Legislativo sobre el manejo que la administración pública realizaba de los ingresos y egresos que autorizan las leyes y presupuestos y, en general, sobre la actividad financiera y patrimonial.

En la actualidad, las actividades internas de control y evaluación de la administración pública, están a cargo de: la SHCP, que continúa con su responsabilidad en materia financiera; la SFP, que persigue el fortalecimiento de los sistemas de gestión y control, el combate a la corrupción y el fomento de la transparencia de la administración pública federal; y la ASF de la Cámara de Diputados, que fiscaliza la cuenta pública. Sin embargo, la decisión del Congreso para que se realicen evaluaciones externas a los programas que reciben recursos fiscales, iniciada en el año 2001,<sup>29</sup> está provocando un enorme aumento en la cantidad de estos trabajos, realizados mediante la contratación de consultorías privadas de evaluación externa, que enfrentan nuevas restricciones técnicas y financieras, mismas que serán analizadas en la segunda parte de este libro. El Banco Mundial reconoce los esfuerzos realizados por el gobierno para mejorar especialmente el monitoreo y la evaluación de los programas sociales (Banco Mundial, 2004a).

Compartimos la idea de que “La evaluación gubernamental sólo es posible en un sistema político donde el debate parlamentario está enriquecido por posiciones partidistas distintas del partido que tiene el poder, pero sobre todo, por una opinión pública concientizada, cuyo papel en los intercambios gobierno-sociedad sea tomado seriamente en cuenta...” (Pardo, 2000: 250). Contamos con la primera condición, pero aún falta bastante para la segunda. Por otra parte, la evaluación realizada por el Poder Legislativo sólo resulta garantía de objetividad si se crean o fortalecen los apoyos técnicos que le permitan realmente analizar la información que reciben.

Obviamente, falta desarrollar un mecanismo efectivo de participación ciudadana en estos procesos, que demuestre que la evaluación no sólo responde a la escasez de recursos y racionalización del gasto, sino que el gobierno se preocupa también por impulsar la transparencia y el desarrollo de la democracia. Incluso las organizaciones sociales y civiles mexicanas han estado impulsadas por su voluntad de contribuir a la resolución de problemas y han considerado durante mucho tiempo a la evaluación como una forma de

<sup>29</sup> El BM (2004b) señala que la decisión se tomó en 1999 para aplicarse al siguiente, pero en nuestra revisión del PEF 2000 no encontramos esa referencia, que en cambio aparece en el PEF 2001.



distraer sus recursos financieros y humanos en actividades no sustantivas. A esto se agrega el rechazo que tradicionalmente generan las evaluaciones provenientes de las fundaciones y organismos que aportan su financiamiento, lo que provocó que se las concibiera más como un control externo no deseado, que como un proceso propio que puede ayudar al aprendizaje de los participantes y la comunidad en su conjunto. Dichas razones han contribuido para que hasta fechas muy recientes se preocuparan muy poco de capacitarse en esta materia y, en todo caso, prefirieran las evaluaciones cualitativas, que si bien permiten profundizar en casos concretos no nos ofrecen una visión agregada de conjunto.

Merece un último comentario la poca preocupación universitaria y de otras instituciones educativas y de investigación por el tema. Si bien fueron éstas las que iniciaron la realización de evaluaciones independientes, a la fecha no existen programas especialmente enfocados al desarrollo de esta incipiente profesión<sup>30</sup> y puede constatarse el reducido número de tesis (29) vinculadas al tema que se ha podido localizar.<sup>31</sup> Además, existe una enorme carencia de material didáctico propio: la inmensa mayoría de los libros y revistas que exponen métodos de evaluación y presentan experiencias de aplicación, provienen especialmente de los EUA y, en mucho menor medida, de Europa o Canadá; respondiendo a contextos muy diferentes de los nuestros. Pocas fueron las evaluaciones hechas desde la academia antes del año 2000 y, a nivel institucional, el seguimiento fue mínimo, no difundido y a menudo incluso se limitó su acceso al investigador.

## Comentarios finales

El capítulo nos permite comprender las dificultades técnicas y políticas encontradas por AL para impulsar procesos de evaluación, que como se vio en el capítulo previo, no son exclusivas de esta región ni producto, al menos único, del nivel de desarrollo; así como la contribución de los organismos internacionales al desarrollo de la misma.

Colombia y Costa Rica se adelantan, a inicios de los noventa, en la implantación de un modelo de evaluación de programas, con la intención

<sup>30</sup> El ITAM es el único que cuenta con un centro especializado en materia de evaluación, pero se trata de un enfoque *ex ante*, especialmente dirigido a proyecto de desarrollo regional.

<sup>31</sup> En una revisión de títulos efectuada en 2003 a través de las páginas web de instituciones como el ITAM, la Universidad Iberoamericana (UIA), el Instituto Tecnológico de Monterrey, la Universidad Nacional Autónoma de México (UNAM), la Universidad Autónoma Metropolitana (UAM) y la Universidad Panamericana, se encontraron: dos tesis de finales de los setenta en la UI y UNAM; ninguna en los ochenta; cuatro en la UI, cuatro en la UNAM y seis en la UAM en los noventa; y seis del ITAM y siete de la UAM posteriores al año 2000.

de reforzar las decisiones de políticas ligadas al Poder Ejecutivo; mientras Chile y Uruguay buscan resultados medibles para orientar las decisiones presupuestales de sus poderes legislativos.

El caso de México aparece como particularmente lejano de EUA y Reino Unido debido a la importancia que en ellos adquiere el Poder Legislativo en la contratación de evaluaciones tendientes a vigilar el desempeño del Ejecutivo y en la amplia disponibilidad de recursos financieros y profesionales para su realización. De los EUA lo separan además, unas cinco décadas en la institucionalización del sistema de evaluación de programas.

En cambio, México se muestra más cercano a la situación descrita en Francia antes del año 2000, caracterizada por la tradición de centralismo y la resistencia a la evaluación, el Poder Legislativo al margen del proceso, retraso en los procesos de rendición de cuentas, falta de desarrollos metodológicos propios, medición de resultados más que de impactos, y limitada utilización de los resultados. No obstante, el caso francés se distancia de la evaluación basada en la experimentación para buscar un modelo propio, que responda al incremento de la democratización y el pluralismo de actores, lo que no sucede en el caso mexicano.

En el contexto latinoamericano, México avanza en la generalización de la evaluación de sus programas sociales (el resto lo hace en forma selectiva) y en el esfuerzo aplicado en el caso particular del programa de Desarrollo Humano Oportunidades. Como veremos en la segunda parte del libro, comparte con los demás países de la región los problemas de sus sistemas de información, medición de impactos, capacitación, aprendizaje social y utilización de los resultados, así como el apoyo recibido de organismos internacionales como el BM o el BID. En cambio, no cuenta aún con una cultura de la evaluación orientada a resultados como la mencionada en Colombia, una difusión de los informes que mantenga informada a la ciudadanía como estaría sucediendo en el caso de Costa Rica y procesos de capacitación apoyados con la preparación de guías de evaluación como se hizo en Reino Unido, Francia y los cuatro países de AL mencionados.

La observación realizada en México muestra que la evaluación actual de políticas y programas públicos cuenta con poca información publicada hasta el presente sobre los procesos de evaluación emprendidos, en particular: 1. los métodos y técnicas aplicados; 2. la calidad de los resultados alcanzados; y 3. el nivel de utilización de dichos resultados en los procesos decisorios. Se trata entonces de responder a preguntas como: ¿por qué, qué y cómo se evalúa?, ¿quien lo hace?, ¿cuánto cuesta?, ¿qué problemas enfrenta?, ¿para qué sirve? Resulta urgente modificar este panorama para promover la participa-

ción ciudadana en los procesos de políticas públicas, contribuir a un proceso de aprendizaje, lograr más transparencia en las decisiones sobre el uso de recursos, eficacia en su aplicación a los objetivos perseguidos y pluralidad para contribuir a la democratización del país.

Para responder a las preguntas previas, en el siguiente capítulo se analizarán los aportes que puede realizar la aplicación de metaevaluaciones a las evaluaciones realizadas.

## Capítulo 6

# Metaevaluación de las evaluaciones

### Introducción

ESTE CAPÍTULO pretende revisar: 1. las razones que limitan el desarrollo de la evaluación; 2. los avances teóricos que se han producido en materia de metaevaluación, entendida como la “evaluación de las evaluaciones”, con objeto de comprender y mejorar la calidad del trabajo de evaluación realizado; 3. algunas de las limitadas experiencias de metaevaluación actualmente disponibles; 4. formular una propuesta aplicable al contexto mexicano en el área de desarrollo social, misma que se desarrollará en la segunda parte de este libro.

### Dificultades para implantar procesos de evaluación

Si bien en los países escandinavos, y otros como Estados Unidos o Canadá, la idiosincrasia, el nivel de desarrollo y/o el sistema político, han permitido mayores o más rápidos avances, como hemos podido ver también existen naciones que, aun gozando de altos grados de desarrollo como Alemania, Reino Unido, Canadá o Francia, han enfrentado importantes resistencias para su aplicación.

Generalmente, las dificultades aumentan en países menos desarrollados como los latinoamericanos y se agudizan en casos como el mexicano, en que su historia política ha estado signada por un agudo presidencialismo y, hasta hace poco tiempo, por la existencia de un partido dominante, mismos que estuvieron lejos de desarrollar la rendición de cuentas a la ciudadanía y la cultura de la evaluación.

Como ya mencionamos, Nioche (1982) identificaba, en Francia, tres tipos fundamentales de problemas ligados al atraso relativo de una cultura de la evaluación en su país, que aparecen como muy pertinentes para el contexto mexicano: sociopolíticos, administrativos y metodológicos.

Los primeros son considerados los más importantes y, al mismo tiempo, los más difíciles de cambiar porque están enraizados en las culturas y tradi-

ciones de los pueblos y sus gobiernos. La evaluación se vuelve amenazante en este contexto porque introduce la duda y la posibilidad de cuestionamiento entre los altos funcionarios, acostumbrados a comportamientos informales. Así, la demanda de evaluación no ha sido muy fuerte hasta fechas muy recientes, pero tampoco la oferta de trabajos. Otros factores que contribuyen a la explicación son: el débil rol del Poder Legislativo como evaluador de la acción gubernamental, el fuerte control de los medios de comunicación masivos por el Ejecutivo, el cuasimonopolio de la información económica y social por la administración pública y la ausencia de factores facilitadores, los que, por el contrario, han estado presentes en la experiencia de EUA, como su federalismo y tradición liberal (no intervencionista).

Los obstáculos administrativos incluyen:

1. la debilidad de la gestión por programas, que resulta el más evidente;
2. la organización del trabajo: el diseño de planes, programas y presupuestos con objetivos muy vagos y ausencia de metas cuantificadas que faciliten su evaluación; y una organización que no define con claridad una entidad administrativa encargada de la realización de una evaluación integral de los programas públicos;
3. los frenos de naturaleza financiera: la falta de recursos destinados a esta actividad, poco legitimada por su discutida aportación al bienestar social; y
4. el rechazo de algunos funcionarios a una evaluación que consideran inútil o excesivamente restrictiva, dados los controles ya existentes.

Los problemas metodológicos tampoco pueden soslayarse. Como ya hemos visto, todas las técnicas empleadas en evaluación enfrentan serios cuestionamientos que disminuyen su credibilidad.

Para Nioche, las evaluaciones comienzan a desarrollarse en ámbitos caracterizados por un fuerte grado de incertidumbre o un elevado potencial de conflicto. Por otra parte, la evaluación sólo tiene sentido en una sociedad abierta a la libre expresión de los ciudadanos sobre la acción gubernamental, permitiendo una pluralidad de análisis, limitada por dos problemas relevantes: el acceso a la información y la desigualdad de los actores sociales en materia de recursos para realizar sus análisis.

Sin embargo, el análisis de políticas puede evitar dos grandes errores, generalmente presentes en la gestión pública, si se desarrolla en un contexto estructural y político que asegure la libertad de información y el pluralismo de los trabajos: considerar que todo es político y no utilizar las herramientas disponibles para

comprender y mejorar la acción pública; o bien, querer utilizar herramientas técnicas para resolver problemas que son realmente de índole política.

Cuatro años después de realizado el artículo anterior, Deleau *et al.* (1986) percibían ya una atención creciente al tema y ofrecían cinco propuestas para lograr el desarrollo de la evaluación de políticas públicas en Francia: transformar a las políticas públicas en algo evaluable, organizar las competencias, reforzar la evaluación de las políticas públicas en el seno de la administración, aumentar la capacidad de evaluación externa de la administración, facilitar su difusión en la sociedad, e identificar los temas prioritarios.

Para Monnier (1995), las condiciones que facilitan el desarrollo de la evaluación son:

- una relativa independencia de los organismos de evaluación, que permita una crítica responsable y un comportamiento profesional ético;
- una capacidad multidisciplinaria de los evaluadores;
- un acceso suficiente a las fuentes de información administrativa; y
- una pluralidad de órganos de evaluación que permita la emulación entre expertos y su empleo por los decisores como fundamento de sus negociaciones.

Muchas veces el resultado de una evaluación se “archiva” sin ser tomado en cuenta, a pesar de todos los cuidados que se hayan puesto en su preparación. Las razones pueden ser muchas: poca motivación de quienes solicitaron la evaluación, afectación de intereses, lenguaje poco accesible empleado por el evaluador, informe entregado en fecha poco oportuna, etcétera. A menudo, la evaluación entrega complejos informes donde los solicitantes buscaban respuestas sencillas, lo que crea registros distintos que obstaculizan la capitalización de las evaluaciones. En particular, se resalta la faceta política y psicológica de la evaluación que, al emitir un juicio sobre las finalidades, la eficacia y la ética de una acción pública, puede “hacer potencialmente vulnerables a los responsables de las políticas y las organizaciones públicas que son quienes contratan la evaluación” (Bañón, 2003: xxv).

Al respecto existe un interesante estudio sobre las representaciones, discursos y experiencias de quienes contratan una evaluación, resaltando su ambivalencia: por un lado manifiestan la importancia de su realización y por otro se perciben temores y aprehensión durante su aplicación (Martin y Ruffiot, 2001). La retórica del discurso y la institucionalización formal de la evaluación no garantizan el cambio de actitudes y conductas. En el mismo sentido se señala que: “La primera exigencia para que la evaluación sea

posible y eficaz es que sea querida... Hay una evidente contradicción entre un sentimiento generalizado que define a la evaluación como indispensable y las reticencias a su concretización” (Duran, 1999: 188, traducción libre). La explicación del empantanamiento de la evaluación está tanto en “...las expectativas que genera como los temores que despierta, y al mismo tiempo, las dificultades que implica su operacionalización” (Duran y Monnier, 1992, traducción libre). Un análisis psicológico en profundidad (exposición de los evaluados a diversas películas y análisis de sus reacciones) habría dado resultados tan negativos en cuanto a lo que la evaluación simboliza, también en el contexto francés, que nunca fue publicado.

Duran afirma que el conocimiento generado por la evaluación debe ser pertinente para quienes lo necesitan y susceptible de ser asimilado por ellos. El lenguaje habitual de los administradores es el de la gestión, no el de sus consecuencias lejanas. Además la evaluación cuestiona su acción y pone en peligro la existencia misma de la organización o programa evaluado. “Las ideas no deben su eficacia sólo a su fuerza lógica, necesitan relacionarse con el interés práctico de los actores y no deben ser demasiado extrañas a su experiencia” (Duran, 2002: IV, traducción libre). Finalmente, no hay que olvidar que la evaluación puede servir para redistribuir las relaciones de poder en la administración pública. Mientras no interese a sus usuarios potenciales, no tendrá impacto; evaluación y gestión necesitan relacionarse.

## Conceptos e instrumentos de metaevaluación

Definimos provisoriamente a la metaevaluación<sup>32</sup> como un tipo particular de evaluación que enjuicia la calidad y el mérito de lo que hacen los evaluadores profesionales, persigue el perfeccionamiento de sus trabajos y promueve una mayor comprensión de los mismos.

A diferencia de las evaluaciones cuyo objeto de trabajo está constituido por las políticas o los programas aplicados, en el caso de la metaevaluación se toma como objeto de estudio a las evaluaciones realizadas, particularmente sus informes finales. En consecuencia, la metaevaluación se caracteriza por una mayor complejidad, determinada por el número y la naturaleza de las variables a considerar, y también profundidad, definida por sus propósitos y su objeto de trabajo (Díaz, 2002).

<sup>32</sup> Es posible encontrar otro empleo de la palabra metaevaluación: en el caso de Cunill y Ospina (2003: 444) es empleada para referirse a la evaluación a nivel “macro”, de las políticas públicas, por oposición de las efectuadas en los niveles “meso” referido a organizaciones y programas públicos, y del “micro” de desempeño de los funcionarios. También se encuentra este sentido en el trabajo de Sanín (1999) en la CEPAL.

No ha sido frecuente que los principales autores que exponen la teoría de la evaluación la mencionen, tampoco se dispone de modelos desarrollados *ex profeso* ni de una epistemología y metodología para su construcción (Díaz y Rosales, 2003). Sus mayores desarrollos se localizan en el campo educativo y las primeras referencias que todos los estudiosos de la metaevaluación reconocen son los trabajos realizados por Michael Scriven (1967, 1969 y 1976), a quien se atribuye también la paternidad del nombre. En una presentación posterior del tema, Scriven se refirió a la misma en los términos siguientes: “...la evaluación de las evaluaciones –indirectamente la evaluación de los evaluadores– representa una obligación ética y científica, cuando se trata del bienestar de los demás” (Scriven, 1991: 228).

El autor subraya así que esta fase, usualmente terminal, del trabajo de evaluación responde a la responsabilidad exigible a todas las actividades humanas, misma que adquiere especial importancia en el caso de la evaluación, por lo que la metaevaluación representa el imperativo profesional de la disciplina, imprescindible para controlar el sesgo del evaluador (Scriven, 1976).

House (1980, citado en Díaz, 2002) dedica al tema una sección de su libro titulado *Evaluación, ética y poder*. Fundamenta el trabajo del metaevaluador en la percepción selectiva “de las disonancias que se presentan en la evaluación y en su informe”, misma que podría analizarse a través de dos enfoques: las opciones “duras” basadas en la cuantificación de variables, relaciones de causalidad, criterios de eficacia y eficiencia, etcétera; y las opciones “suaves” con gran cantidad de variables cualitativas interactivas implicadas en un trabajo hermenéutico de descripción, comprensión e interpretación de textos. Las primeras opciones podemos vincularlas a definiciones de metaevaluación de contenido práctico y operativo; mientras que las segundas se asocian a la búsqueda más profunda de sentido y explicación de los procesos evaluativos.

House se inclina tanto a la utilización de ejercicios sistemáticos de hermenéutica aplicados al texto (búsqueda de relaciones de identidad, contradicción, reciprocidad, transitividad, etcétera) como al empleo de técnicas cualitativas de manejo de datos (análisis, síntesis, comparación, causalidad, etcétera).

Schwandt y Halpern (1988) proponen dos modelos de metaevaluación. El primero se trata de la revisión por el metaevaluador, solo o conjuntamente con el evaluador del programa, de las etapas desarrolladas en el proceso, entre las que resaltan las de preparación del trabajo: análisis de su factibilidad, negociación del contrato, planeación actividades, etcétera. El segundo, se refiere al proceso mismo de investigación y consta de una lista de comprobación para garantizar las condiciones de confiabilidad y validez de los resultados.



La posición de estos autores nos parece de particular interés por dos asuntos que plantean, a diferencia de los demás, y que se relacionan con la metaevaluación realizada en la presente tesis:

1. la posibilidad de que el metaevaluador no se encierre solo a realizar su trabajo, sino que pueda interactuar con los evaluadores; y
2. la necesidad de analizar, no únicamente el proceso y los métodos de evaluación aplicados, sino también algunos aspectos previos y posteriores, vinculados al funcionamiento del sistema o contexto en que el trabajo se inscribe.

Scriven (1991) propone la aplicación de su Key Evaluation Checklist, compuesta de las siguientes 13 dimensiones o aspectos a revisar, como instrumento de evaluación que también puede ser usado en la metaevaluación:

1. Descripción: diseño, equipo evaluador, calendario del estudio, fuentes de información, etcétera.
2. Antecedentes y contexto: responsable de la solicitud de evaluación, razones, evaluaciones previas, etcétera.
3. Consumidor: clientes y destinatarios efectivos y potenciales.
4. Recursos: insumos que pueden limitar el diseño y ejecución de la evaluación.
5. Valores: criterios de mérito en función de las necesidades del cliente y las normas profesionales.
6. Proceso: desarrollo de la evaluación y problemas encontrados.
7. Resultados: informes y sus efectos previsibles por el evaluador así como impactos del proceso de evaluación.
8. Costos: monetarios y no monetarios, así como sus costos alternativos.
9. Comparaciones: utilización de juicios validados, el azar, etcétera.
10. Posibilidades de generalización: posibles usos en diseños futuros, otros programas, otros lugares, etcétera.
11. Relevancia: síntesis de todo lo anterior que permite valorar el mérito de la evaluación realizada.
12. Recomendaciones: se pueden o no hacer en torno a la continuidad del programa y sus posibles ajustes.
13. Informe del evaluador.

Stufflebeam y Shinkfield (1987) se acercan al concepto de metaevaluación definido por Schwandt y Halpern (1988) para incluir la valoración "...de

todo lo que representa y hace la profesión”, con objeto de garantizar que la evaluación cumpla con estándares de adecuación técnica, costo-efectividad y utilidad, incluyendo la planeación de la evaluación, la investigación evaluativa, el desarrollo organizativo, los servicios de evaluación y la utilización de sus resultados. Retoman la lista de Scriven y otra similar preparada por Stufflebeam en 1974 en la Western Michigan University y elaboran una versión depurada que diferencia entre el cliente que encarga la evaluación y sus consumidores, y que agrega nuevos elementos como la función que cumple el evaluado, el sistema que utiliza para distribuir sus servicios, el significado o síntesis de todo lo analizado y la metaevaluación que puede incorporarse al propio informe de evaluación.

La metaevaluación también ha sido reconocida por especialistas más recientes como Monnier (1995), Patton (1997), Weiss (1998), Bañón (2003), etcétera, y por organizaciones como el Joint Committee on Standards for Educational Evaluation (1994).

Díaz (2002) cita algunos trabajos adicionales<sup>33</sup> entre quienes resalta Santos Guerra (1999), quien además de elaborar un inventario sistemático de los elementos a comprobar (necesarios pero no suficientes para evaluar el trabajo realizado), elaboró preguntas relevantes para los grandes momentos de la evaluación:

- constitución: origen de la evaluación, negociación de sus términos, etcétera;
- justificación: rigor metodológico, validez, tratamiento de la información, etcétera;
- presentación: estructura lógica del informe, extensión, coherencia, claridad, etcétera;
- negociación: rapidez, intermediación, jerarquización en la entrega del informe;
- valoración: motivaciones, intenciones, presiones, actitudes de los participantes; y

<sup>33</sup>Entre ellos incluye cuatro artículos relevantes y una bibliografía sobre el tema. Los primeros son: Cook y Gruder (febrero de 1978), “Meta-evaluation Research”, en *Evaluation quarterly* 2(1); Reineke y Wech (1986), “Client-centered meta-evaluation”, en *Evaluation practice*; Santos Guerra (1999), “Metaevaluación: rigor, mejora, ética y aprendizaje”, en Bonifacio Jiménez Jiménez, *Evaluación de programas, centros y profesores*, Madrid, Síntesis citado en Díaz, 2002), y J.F. Angulo (1993), “Metaevaluación del programa de educación de diabéticos”, en Sorriquer *et al.*, citado por Bonifacio Jiménez, *Evaluación de programas, centros y profesores*, Madrid, Síntesis.

La bibliografía corresponde a:

-Patricia Rogers, (1997, comp.), *Metaevaluation bibliography*, Program for Public Sector Evaluation, Royal Melbourne Institute of Technology.

–difusión: distribución de la información y resultados, demoras, adulteraciones, etcétera.

Compartimos con Díaz la idea de que la metaevaluación puede enfocarse en forma restringida a los datos y hechos que constituyen el contenido técnico de la evaluación (procesos y resultados), que la mayoría de los autores señalan como su contenido, y operacionalizarse mediante listas de comprobación. Pero también puede trascenderlos para tomar en cuenta todos los factores que la afectan. Para Díaz dichos factores incluyen: evaluadores, modelos, categorías, fundamentos metodológicos y epistemológicos, paradigmas científicos, ideología, filosofía, usos de lenguaje, codificación, conducta del evaluador, valores en juego, etcétera.

Sin embargo, el autor reconoce que generalmente no se cuenta con información de este cúmulo de aspectos, ya que la metaevaluación se realiza con base en los informes oficiales del evaluador, por lo que recurre a proponer el método hermenéutico para desentrañar la ideología, la concepción del mundo y la epistemología que fundamentan los modelos y categorías orientadores del proceso evaluativo. Aclara que dicho método "...consiste en repetir sistemática y cíclicamente, a lo largo de todo el proceso, el ejercicio representado por las cuatro funciones: análisis estructural, interpretación, comentario y crítica" (Díaz, 2002: 329).

Santos Guerra refiere la necesidad de interrogarse sobre la naturaleza de la iniciativa, la finalidad de los procesos, la utilización de sus resultados, el acceso democrático a los informes, etcétera; conjunto de factores hacia el que parecían apuntar las anteriores propuestas de Stufflebeam y Shinkfield y de Schwandt y Halpern.

Para García Sánchez (2003 y 2004), trascender las características más evidentes de las evaluaciones implica desentrañar sus causas mediante el estudio de las decisiones de los actores y sus motivaciones, recuperando las contribuciones pluralistas, lo que coadyuva a la comprensión de la forma en que se distribuye el poder en estos procesos. Propone dar seguimiento a las decisiones de mayor relevancia en el proceso de evaluación:

- selección del equipo evaluador y contrato de evaluación;
- enfoque metodológico;
- selección de actores participantes;
- objeto y objetivos de la evaluación;
- criterios de valor;
- técnicas de recolección y análisis de datos;
- contenidos y estilo del informe de evaluación;

- difusión y/o publicación del informe; y
- utilización del proceso y los resultados de la evaluación.

En este trabajo consideramos que este aditamento se conforma por factores de carácter organizacional, que es posible reconstruir echando mano de la técnica de entrevista aplicada a informantes-clave, como complemento del análisis de informes.

Tratando de esquematizar la revisión de conceptos e instrumentos realizada, identificamos:

- metaevaluaciones dirigidas exclusivamente al análisis del informe final; otras orientadas a la revisión de los procesos técnico-metodológicos implicados en la investigación; o, aquellas que se enfocan a los anteriores en el marco de un contexto político y una determinada estructura y funcionamiento organizacional;
- metaevaluaciones con enfoque racional, cuantitativo y de perspectiva positivista basado en el análisis estadístico de datos; con enfoque fenomenológico, holista, poco estructurado y cualitativo, fundamentado en la hermenéutica y el estudio de los aspectos políticos, organizativos o éticos; o, con enfoque mixto; y
- metaevaluaciones que utilizan diversas listas de indicadores o aspectos a revisar, propuestas por distintos expositores del tema.

Por otra parte, Bustelo (2001 y 2002) plantea que existen autores y organizaciones que confunden el concepto de metaevaluación con otros como la síntesis evaluativa y el metaanálisis. Se trata del caso de la Evaluation Research Society que la considera uno de los seis tipos de evaluación que reconoce y la define como una forma de volver a analizar los resultados de una o varias evaluaciones para determinar lo que se ha aprendido en torno a esa política o programa. Vedung (1997) incluye en el concepto de metaevaluación la síntesis de resultados de diferentes estudios de evaluación y Ballart (1992) expresa: "...consiste en el análisis del resultado de varias evaluaciones, y por otro lado, en la revisión de la más amplia investigación desarrollada en un cierto periodo".

Para la autora los tres conceptos que se mezclan son diferentes:

- la síntesis evaluativa resume los resultados de múltiples evaluaciones efectuadas a políticas o programas similares, y sólo se interesa en la

- metodología como la base que permitió llegar a los resultados (no analiza los procesos);
- el metaanálisis consiste en un procedimiento estadístico para comparar los hallazgos de diferentes estudios que confluyen en un mismo objeto, identificar tendencias y realizar generalizaciones, que ha sido aplicado a las evaluaciones (Glass, 1976); y
  - la metaevaluación no se enfoca a los resultados de las evaluaciones sino a los procesos implicados en la realización de una o varias de ellas y, cuando incluye a los primeros, no es para sintetizarlos sino para revisar si cuentan con evidencia suficiente, si la metodología ha sido adecuada, si las interpretaciones, juicios y recomendaciones fueron lógicamente diseñados, etcétera. Cumple dos funciones básicas: la de descripción, análisis y valoración de los procesos de evaluación, y la de seguimiento, fomento y control de la calidad de las evaluaciones, mismos que también permiten asegurar la credibilidad de la evaluación en contextos muy politizados.

Bustelo agrega que la literatura muestra una tendencia a metaevaluar una única evaluación,<sup>34</sup> ya que son raros los procesos de evaluación que se analizan y comparan simultáneamente debido a que el trabajo se guía más por la lógica y tiempos de la acción pública que por los de la construcción del conocimiento.

En conclusión, dentro del presente libro se comparte, pero se amplía la definición presentada al inicio ya que se considera que la comprensión y el perfeccionamiento de la evaluación debe llevarnos a enjuiciar el valor y el mérito no sólo de los trabajos de los evaluadores profesionales, sino el de todo el sistema en que se organiza y desarrolla el anterior, incluyendo el papel jugado por otros actores o instancias que inciden en la calidad del proceso desarrollado y del informe final logrado.

Esto implica no reducir la metaevaluación al estudio de los métodos y procesos aplicados por el evaluador, sino recuperar fundamentalmente las propuestas que desde la teoría de la organización y el enfoque pluralista sugieren tomar en cuenta el contexto inmediato en que la evaluación se desarrolla, considerada ésta como un proceso social complejo. Por otra parte, significa que ninguna disciplina ni enfoque puede pretender ser el único adecuado para realizar una metaevaluación, por lo que su análisis debe ser de tipo multidisciplinario y mixto.

<sup>34</sup>Sin embargo, no es lo que encontramos en nuestra investigación.

## Criterios de metaevaluación

Aún son muy escasas las propuestas de criterios específicos para la metaevaluación.

House (1980, citado en Díaz, 2002) propuso criterios muy generales, poco operativos, que recuperan en forma amplia los procesos relacionados con la evaluación, tales como veracidad, autonomía, credibilidad, participación, equidad, poder, objetividad, etcétera, los cuales considera que deben coadyuvar al logro de una mayor justicia social.

Patton (1997) señala que el debate entre cuantitativistas y cualitativistas sirvió para mostrar que no hay criterios universalmente válidos en todos los casos; sin embargo, también plantea que existen algunos de carácter general que pueden utilizarse como referencia: conveniencia, utilidad, factibilidad, credibilidad y relevancia.

Scriven (citado en García: [www.ucm.es](http://www.ucm.es)) elabora, a partir de diferentes listas, una síntesis de aspectos específicos denominada Meta-Evaluation Checklist, que incluye las características deseables a verificar en una metaevaluación: estudio comprensivo y comprensible, creíble, ético, factible, suficiente, oportuno, seguro, válido, sensible desde el punto de vista político, que explicita los estándares de valor utilizados y su justificación, y costo-eficaz.

En cambio, Stufflebeam (1974, citado en Centeno, 2002) recupera los principales criterios metodológicos y epistemológicos, en relación exclusiva con los procesos de investigación: validez interna, validez externa, confiabilidad, objetividad, relevancia, importancia, alcance, credibilidad, tiempo, difusión, y relación costo-efecto.

Schwandt y Halpern (1988) también identifican criterios específicos de calidad de una metaevaluación, sin tomar en cuenta su contexto: el rigor de la metodología utilizada y su capacidad de convencimiento basada en la bondad de los resultados hallados en la evaluación.

La primera se refiere a los procedimientos utilizados para conocer el objeto de estudio (evaluación): coherencia epistemológica, metodológica y tecnológica, adecuación a su contexto, relevancia para la política o programa evaluado, definición de criterios preestablecidos de evaluación, técnicas de recopilación, análisis de datos, participación de actores relevantes y relación costo-beneficio de la evaluación.

La segunda trata de conseguir la coherencia lógica y el señalamiento independiente y la vinculación de los cuatro niveles definidos por Patton (1997), en que cada nivel se basa en el anterior: datos, interpretaciones, valoraciones y recomendaciones.

Schwandt y Halpern definen los siguientes criterios para valorar los datos: relevancia, confiabilidad, suficiencia, representatividad y temporalidad. Las interpretaciones y explicaciones obtenidas a partir de los datos que cumplen los criterios anteriores serán adecuadas si su análisis es correcto y suficiente. Las valoraciones deben hacerse en función de los criterios preestablecidos en acuerdo con los diferentes agentes de la evaluación. Las recomendaciones deben basarse en los niveles previos y considerar el contexto en que se desarrolló la acción evaluada; ser realistas, claras y se recomienda que se acompañen de una estimación de sus costos para facilitar su posible implementación.

Finalmente, las normas más utilizadas como referencia para la metaevaluación son las del Joint Committee, que aparecen ordenadas en cuatro grupos: utilidad, factibilidad, corrección/legitimidad, y precisión (véase anexo 1). Las mismas son reconocidas por su especificidad, claridad, orden y estructuración, pero cada metaevaluación puede definir sus propios criterios (Bustelo, 2001).

García (en *Diccionario Crítico de Ciencias Sociales*: [www.ucm.es](http://www.ucm.es)) menciona que lo importante es valorar la capacidad explicativa de la investigación evaluativa, por lo que los criterios se deben determinar en función de los objetivos de la investigación y de la perspectiva del analista.

En conclusión, existe una amplia gama de criterios a seleccionar en cada caso, mismos que incluyen:

- los de carácter general, relativos a los procesos de investigación: objetividad, validez, confiabilidad, representatividad, suficiencia, veracidad, relevancia, relación costo-beneficio, etcétera;
- los específicos de la metaevaluación: autonomía, credibilidad, utilidad, factibilidad, oportunidad, participación, legitimidad, etcétera.

### Tipos de metaevaluación

A las metaevaluaciones pueden aplicarse todas las clasificaciones de evaluaciones, ya revisadas.

Los enfoques *cuantitativistas* enfatizan el análisis de los aspectos metodológicos y, en casos extremos, identifican como modelo de calidad del trabajo el empleo de diseños experimentales (Campbell y Stanley, 1966).

Las corrientes *cuantitativistas* enfatizan los aspectos éticos y políticos de la investigación, a veces incluso con descuido de las cuestiones metodológicas (House, 1980, citado en Díaz, 2002 ; Scriven, 1991).

Las *metaevaluaciones de diseño* deben analizar el contexto evaluativo, la claridad de sus objetivos, la inclusión de las variables relevantes, la justificación del esfuerzo, su costo, el tiempo disponible, su contribución potencial a la mejoría del programa, la participación de los actores involucrados, la disponibilidad de la información necesaria, la previsión de técnicas adecuadas de recolección y análisis de datos, de presentación y difusión de informes, etcétera.

Las *metaevaluaciones de proceso* analizan la puesta en práctica de la evaluación antes diseñada: ajuste al calendario, disponibilidad oportuna y uso eficiente de recursos, calidad del equipo evaluador, aplicación de los métodos y técnicas previstas, problemas encontrados, etcétera.

Algunos autores consideran como resultados de la *metaevaluación* tanto sus *productos* como sus *impactos*. En la primera, se revisa la calidad de los productos típicos de una evaluación, que incluyen datos, interpretaciones, valoraciones y recomendaciones, así como la forma en que son recuperados en los informes. También se analiza si estos se entregaron a tiempo y se difundieron.

La segunda se refiere a los efectos que provoca la evaluación, su utilización efectiva: modificaciones al programa, contribuciones a la política en que se enmarca, procesos de aprendizaje, etcétera (Bustelo, 2001). Weiss (1998) define cuatro posibles impactos o usos de la evaluación: instrumental (para su uso en la toma de decisiones), conceptual (para cambiar la comprensión de lo que el plan o programa hace), persuasivo (cuando moviliza apoyos para efectuar los cambios requeridos por el programa), y de influencia en redes de políticas, programas, evaluadores, etcétera, ajenos al programa evaluado.

## Experiencias de metaevaluación

Bustelo (2001 y 2004) plantea que, a pesar del desarrollo alcanzado por la evaluación en Europa, aún hay muy poca investigación que sistematice y analice estas experiencias (metaevaluación). En los países en que la cultura y la práctica de la evaluación están más desarrolladas se ha insistido, en los últimos años, en la necesidad de realizar una investigación que aproveche los estudios ya realizados. Algunos antecedentes pueden encontrarse en Reino Unido, Canadá o Bélgica.

En el presente libro se mencionan los trabajos a los que se tuvo acceso: varios realizados en Francia y España, además del canadiense ya mencionado, y la única localizada en AL. Casi todos refieren sus técnicas de recolección de la información, pero ninguno explicita el tratamiento posterior que dio a la misma para arribar a sus conclusiones.



Constituye un primer trabajo de metaevaluación el realizado por Lacasse (1993) en Canadá, referido dentro del segundo apartado de este capítulo. Para su estudio, inició un diagnóstico de la situación, basado en la revisión de 37 informes de evaluación del periodo 1988-1990, complementado con entrevistas confidenciales aplicadas a funcionarios, consultores y clientes de las evaluaciones, con base en una muestra intencional en que seleccionó sólo los trabajos de mayor calidad y dirigidos por personas de reconocida competencia.

Lascoumes (1998) relata una metaevaluación francesa que dio origen a la realización del coloquio “Evaluation pluralista des politiques publiques: enjeux, pratiques et produits”. Su objetivo era enfocarse al conjunto de trabajos colectivos que permitían la materialización de la evaluación: diversidad de prácticas, herramientas, modelos teóricos y metodologías. Especialmente se trató de saber si los evaluadores habían podido conciliar y cómo los objetivos de racionalización y coherencia de la política con los de su legitimación y aceptación, ya que el modelo francés de 1990 conservaba elementos de una evaluación técnica o gerencial, al mismo tiempo que introducía el enfoque pluralista; negaba prioridad al saber de los expertos y abría el proceso a los “profanos implicados”.

Como objeto de observación se seleccionaron seis evaluaciones con diferentes encuadres para averiguar si éstos provocaban consecuencias en los procesos y productos de la evaluación: cuatro evaluaciones presentaban el modelo Cime-CCE,<sup>35</sup> la quinta, aunque también era interministerial, se realizó en un encuadre estrictamente administrativo;<sup>36</sup> y la sexta no fue interministerial ni nacional, se trataba de la evaluación de una política pública local, realizada en un encuadre similar al CIME-CCE, aunque con algunas diferencias significativas.<sup>37</sup> Las cuatro primeras pudieron llegar a conclusiones prudentemente generalizables; las dos restantes hicieron un aporte distinto, sin pretensiones de representatividad de ningún tipo.

Analizaron todos los documentos posibles, producidos por las instancias de evaluación, así como aplicaron entrevistas a sus miembros y a algunos actores externos (solicitantes y encargados de la evaluación).

Dedujeron que la construcción metodológica de cada instancia dependía de cuatro factores:

<sup>35</sup> La informática en el Estado, La inserción de los adolescentes en dificultades, La rehabilitación del hábitat social, y El desarrollo de infraestructura en las zonas húmedas.

<sup>36</sup> La estrategia de formación de trabajadores sociales, realizado por el Ministerio de Asuntos Sociales.

<sup>37</sup> Política pública regional de turismo rural en la región de Bretaña.

1. el contexto político en que se seleccionaba al responsable de la evaluación y al objeto y problemática que ésta cubriría;
2. el tipo de relación que se establecía entre los diferentes grupos e intereses al interior de la instancia de evaluación;
3. el respaldo científico brindado por el CCE o instancia equivalente, que debería apoyar en forma continua el proceso metodológico pero sólo participa al inicio y al final del mismo; y
4. las instancias, que constituyen intelectuales colectivos que se diferencian por su dinámicas internas y sus producciones.

Concluyeron que la evaluación no es un factor de rediseño de las políticas públicas; su impacto es más sensible y profundo como forma de esclarecer problemas y lograr un aprendizaje que podría transformarse en un proceso de cambio al interior de la administración. Otra de sus conclusiones fue que el modelo pluralista francés enfrentaba tres tipos de problemas:

1. la dificultad para superar la diferencias entre la lógica del solicitante, que busca controlar el proceso, con la de un pluralismo real;
2. la tendencia a ocultar los objetivos de la política a evaluar, que no siempre pueden ser descubiertos por la instancia de evaluación; y
3. las consecuencias efectivas de la evaluación, más indirectas que directas, y raramente sostenidas por quienes iniciaron el proceso.

Bustelo (2001), en su trabajo aplicado a la metaevaluación de políticas de igualdad de género en España, analiza 11 planes y programas en el tema. Utiliza técnicas de recopilación y análisis documental; entrevistas abiertas semiestructuradas aplicadas a los responsables de los organismos involucrados; y entrevistas estructuradas, personales y telefónicas, realizadas a los responsables técnicos de la evaluación en los mismos organismos. Señala que existe un claro déficit de valoraciones globales en las evaluaciones aplicadas y que ninguno de los informes revisados en forma detallada contiene un apartado de recomendaciones, lo que influye en la falta de utilidad y de impacto de dichos trabajos.

Lajoudie y Toulemonde (2002) analizaron 11 evaluaciones realizadas en la zona de Limousin, Francia, entre 1994 y 1998, en relación con sus procesos y la calidad de sus informes, y seleccionaron cuatro de ellas para profundizar en la utilidad de las mismas. Los temas revisados incluyeron: calidad de la solicitud de evaluación, conducción de la evaluación, calidad del informe y explotación de sus resultados. Para la revisión de los métodos de evaluación

empleados utilizaron la aplicación de entrevistas directas, por teléfono y por correo electrónico. También revisaron los principales cambios acontecidos en el periodo. Entre sus conclusiones resalta la poca utilización de los resultados de las evaluaciones.

Larrú (2003) realizó una metaevaluación del sistema de evaluaciones de EuropeAid para valorar la calidad y la utilidad de los informes de evaluación externa contratados por su Unidad de Evaluaciones de la Unión Europea. Concluye que existe una gran variedad de tipos de evaluación aplicados (intermedias, *ex post*, sumativas, formativas, etcétera), que las recomendaciones son pertinentes aunque a veces demasiado generales, algunos informes demasiado extensos (más de 100 páginas), se presentan riesgos de minimizar costos a expensas de la calidad, y dudas sobre la influencia real de las recomendaciones efectuadas y su impacto en la asignación de recursos.

Plantea además algunos dilemas entre los que incluye que las evaluaciones se dirigen a quienes las financian y elaboran unilateralmente los términos de referencia, por lo que las mismas ayudan muy poco a los beneficiarios de los programas y nunca enjuician la actuación de los donantes. También identifica futuras líneas de investigación que requerirían la aplicación de encuestas a funcionarios para complementar los datos incluidos en los documentos que analiza (costo de las evaluaciones, términos de referencia, etcétera).

También en España existe un trabajo de metaevaluación, de enfoque cualitativo, efectuado por Ester García (2003 y 2004). Fue aplicado a tres de las evaluaciones más destacadas de programas y proyectos de reforma educativa, realizadas en el periodo 1982-1992. Se basó en el enfoque de actores y trató de encontrar las causas que explican las decisiones adoptadas por los mismos.

Sus principales conclusiones incluyen:

1. los intereses que los actores no mencionan llegan a ser los que guían realmente los objetivos de las evaluaciones;
2. la mayoría de las decisiones fueron tomadas por un número reducido de actores;
3. las decisiones sobre el enfoque metodológico son las de mayor relevancia porque influyen en el resto de las decisiones críticas de la evaluación;
4. las evaluaciones analizadas no cuestionaron los programas y sus premisas;
5. tuvieron un escaso impacto en el perfeccionamiento y reorientación de los programas;

6. no se detectó correlación alguna entre el grado de pluralismo de las evaluaciones y la utilidad de sus resultados, etcétera.

En conclusión, cinco de las seis metaevaluaciones revisadas concluyeron que los resultados de las evaluaciones son muy poco utilizados y el sexto (Larrú) manifiesta dudas al respecto. En el mismo sentido se expresan Díaz y Rosales (2003) con base en las evaluaciones que revisan en el contexto costarricense.

Como ha podido verse hasta el momento, es en el continente europeo donde se han identificado más casos de metaevaluación. En AL sólo se identificó un trabajo de tesis, desarrollado por Centeno (2002) en la Universidad de Costa Rica, que trata de realizar una metaevaluación del Programa Antidrogas y Salud Sexual y Reproductiva para Adolescentes, aplicado en Jamaica, en 1999-2002.

A pesar de que, según Bustelo (2001), existe una tendencia a metaevaluar una sola evaluación, los casos presentados hasta el momento se referían todos a múltiples evaluaciones; en cambio, el trabajo de Centeno, sólo metaevalúa un programa. Si bien el mismo plantea que analizará en forma integral todo el proceso (no queda claro si de evaluación o de gestión del programa) y que le interesan tanto los aspectos cuantitativos como los cualitativos, su trabajo se refiere únicamente en los aspectos técnicos, que son analizados en forma cuantitativa: las estrategias de análisis estadístico utilizadas durante la evaluación para la obtención de datos, su tratamiento e interpretación.

Realiza su trabajo debido a una discrepancia inicial con la forma en que se obtuvieron los resultados de la evaluación, rebasando el concepto de metaevaluación que hemos manejado, ya que no se limita a analizar críticamente los procesos desarrollados, sino que propone y aplica nuevos métodos de tratamiento a los datos. En realidad, realiza una metaevaluación de los aspectos metodológicos y agrega una segunda evaluación del programa (ajena ya al concepto de metaevaluación), a partir de los mismos datos, aplicando nuevos métodos, lo que le permite obtener resultados también diferentes.

Entre las conclusiones de su metaevaluación resalta la ausencia de una valoración inicial de las necesidades de los beneficiarios; una teoría del programa que relacione los insumos y actividades con el resultado esperado; pruebas de validez y confiabilidad del cuestionario aplicado; y utilización de un método de tratamiento ajustado a los datos recogidos.

Uno de los aspectos que las metaevaluaciones permiten identificar son los impactos logrados por las evaluaciones. Lascoumes (1998) los clasifica en

impactos positivos, o sea los que generan una dinámica de cambio, e impactos negativos (que no generan cambios o incluso provocan resistencias).

Entre las consecuencias positivas identifica: consecuencias técnicas o pragmáticas (nuevos conocimientos, integración de los dispersos, dinámicas de capacitación); acciones de reencuadre de la acción pública (clarificación de competencia entre actores, mayor coherencia en la toma de decisiones, adaptación de recursos); consecuencias heurísticas sobre los actores individuales o institucionales implicados, relacionadas con su implicación en el dispositivo de evaluación, por lo tanto, limitada a un pequeño círculo, poco difundible, temporal; y, consecuencias políticas (reformulación del problema, reingreso a la agenda, cambios estratégicos).

Por su parte, las consecuencias negativas pueden incluir: falta de apropiación de los resultados de la evaluación por las instancias políticas y/o administrativas; rechazo de los resultados por su solicitante debido a cambios en el contexto político (más rara), mala interpretación de la solicitud, o conclusiones imposibles de digerir para la élite gobernante; desnaturalización de los datos y las recomendaciones de la evaluación; e, imposibilidad de un actor político que forma parte del proceso de evaluación de introducir cambios en función de los resultados obtenidos debido a la falta de recursos materiales o simbólicos.

### Propuesta de metaevaluación a aplicar en México

En la segunda parte de este libro se propone realizar una metaevaluación de los programas de desarrollo social a cargo de la Sedesol que cubra los objetivos revisados (describir, analizar y valorar los procesos de evaluación y la calidad de sus resultados) y desarrolle los determinantes de carácter organizativo. Para ello analizará los informes finales de evaluación, los procesos de investigación desarrollados y los elementos del contexto que pudieron afectarlos. A partir de estos procederá, con un enfoque multidisciplinario, cuantitativo y cualitativo para:

- a) ofrecer una visión global de cómo se organiza y funciona la evaluación en el sector, en términos formales e informales;
- b) metaevaluar cada uno de los programas evaluados;
- c) realizar una síntesis de las anteriores como metaevaluación transversal del conjunto; y
- d) analizar la evolución vivida por la evaluación de algunos programas sociales seleccionados.

En términos de las clasificaciones básicas revisadas, dicha aplicación se define como una metaevaluación:

- a) sectorial, porque analiza las evaluaciones realizadas en el sector coordinado por la Sedesol;
- b) formativa, ya que realiza su análisis con fines de perfeccionamiento futuro de esta práctica;
- c) *ex post*, debido a que su valoración se realiza después de su aplicación,
- d) externa, por no tener vínculo directo con los operadores de los programas y haber realizado sólo una de las 31 evaluaciones consideradas, de poco peso en el conjunto; y
- e) integral, o sea, de diseño, proceso, resultados e impactos.

Esta metaevaluación utilizará el análisis documental (leyes, decretos, planes, programas, acuerdos, reglas de operación, términos de referencia, informes de evaluación, etcétera), pero también aplicará entrevistas abiertas y semiestructuradas a sus actores más relevantes (funcionarios, evaluadores, congresistas, organizaciones no gubernamentales, etcétera). En cambio, no hará uso de técnicas propias de la hermenéutica como el análisis de textos.

Las evaluaciones serán analizadas mediante un conjunto de indicadores o lista de comprobación que integran elementos de las distintas propuestas de criterios revisadas y los adaptan a nuestro contexto nacional:

- Decisiones y procesos del contexto inmediato a la evaluación de programas:
  - planeación de procesos de evaluación,
  - origen de la decisión de evaluar,
  - instancias intervinientes y su papel formal,
  - funcionamiento real del sistema:
    - relaciones entre las instancias,
    - licitaciones y términos de referencia,
    - mercado de la evaluación,
    - cumplimiento de obligaciones legales,
    - equipos evaluadores y sus móviles,
    - principales problemas administrativos,
  - recursos para evaluar,
  - difusión de los resultados de las evaluaciones, y
  - utilización de resultados en la toma de decisiones.

- Metodología y procesos internos a la evaluación de programas:
  - fuentes de información y análisis de datos,
  - pertinencia y diseño del programa,
  - recursos utilizados por el mismo,
  - análisis de procesos,
  - medición de cobertura,
  - resultados alcanzados,
  - impactos provocados,
  - participación y satisfacción social,
  - utilización de resultados y de recomendaciones previas,
  - valoración global del programa y recomendaciones,
  - escenarios prospectivos,
  - informes finales, y
  - costo de la evaluación.

Estos temas serán valorados con base en los criterios generales de investigación (objetividad, validez, confiabilidad, representatividad, consistencia, etcétera) y con algunos específicos de la metaevaluación como pertinencia, autonomía, credibilidad, utilidad, participación y calidad.

A su vez, para la medición de la calidad se tomarán en cuenta cuatro elementos que se consideran relevantes para el contexto nacional: 1. haber sido realizada por un equipo interdisciplinario, pertinente al tema del programa evaluado; 2. contar con un diseño, recolección y tratamiento de datos rigurosos; 3. cumplir con los requerimientos especificados en el Acuerdo y los Términos de Referencia; y 4. concluir con la explicitación de un juicio valorativo fundamentado sobre el programa.

## Comentarios finales

El capítulo sintetiza las dificultades culturales, sociopolíticas, administrativas y metodológicas que enfrenta el desarrollo de procesos generalizados de evaluación y los aportes que al respecto pueden realizar los trabajos de metaevaluación, dirigidos a enjuiciar el valor y el mérito de los trabajos de evaluación, tanto en términos metodológicos como organizacionales.

Se presentan los limitados avances teóricos, metodológicos y empíricos producidos por distintos países en materia de metaevaluación, y se menciona la localización de un solo trabajo de este tipo en AL, resaltando la poca utilización de sus resultados en todos los ámbitos.

Finalmente, se formula una propuesta de metaevaluación multidisciplinaria, sectorial, *ex post*, sumativa, externa y de diseño, procesos, resultados e impactos, que se desarrollará en la segunda parte del presente libro. La misma nos guiará en la segunda parte de este libro para la realización del balance de los avances logrados en México en materia de evaluación de programas de desarrollo social, permitiéndonos analizar los problemas aún no resueltos y proponer soluciones.



Segunda parte

LA EVALUACIÓN DE POLÍTICAS  
Y PROGRAMAS MEXICANOS  
DE DESARROLLO SOCIAL

## Capítulo 7

# Políticas y programas de desarrollo social

### Introducción

EN EL PRESENTE capítulo se analizan las particularidades de las políticas sociales y su evolución, y se realiza una breve síntesis de los principales programas sociales dirigidos al combate de la pobreza, enfatizando las características de los que han estado vigentes en el sexenio actual.

### Política social y pobreza

Las políticas sociales constituyen un tipo particular de políticas públicas, cuyo objetivo final es el logro del bienestar social mediante estrategias de redistribución de la riqueza y del ingreso para impulsar un mejoramiento continuo de las condiciones de vida de la población. Impulsadas en Europa en el marco del Estado de bienestar al finalizar la Segunda Guerra Mundial, y en menor medida por los latinoamericanos, tendieron a ampliar los derechos sociales y procurar condiciones aceptables de vida para el conjunto de la sociedad.

La política social ha sido una función asumida por el Estado mexicano, con énfasis variables, junto a sus funciones clásicas de justicia, seguridad y orden público.

En los años cuarenta, el proceso de industrialización mexicano se caracterizó por provocar altas tasas de crecimiento económico acompañadas de una mala distribución del ingreso entre el campo y la ciudad y por grupos sociales, lo que generó exclusiones que se irían profundizando con el paso del tiempo. Sin embargo, aun los más pobres vieron aumentar su nivel de ingreso debido a la existencia de un fuerte y constante crecimiento económico en el periodo.

Como ya se mencionó en la revisión de antecedentes, México desarrolló múltiples programas sociales a cargo de diversas organizaciones: SSA, Secretaría de Educación Pública (SEP), Sedesol, Instituto Mexicano del Seguro

Social (IMSS), Instituto de Seguridad y Servicios Sociales para los Trabajadores del Estado (ISSSTE), Instituto Nacional del Fondo para la Vivienda de los Trabajadores (Infonavit), Fondo Nacional de las Habitaciones Populares (Fonhapo), Instituto Nacional Indigenista (INI), Conasupo, Liconsa, Diconsa, Instituto Nacional de Desarrollo Social (Indesol), Comisión Nacional de Zonas Áridas (Conaza), Fondo Nacional de las Artesanías (Fonart), entre otras.

Sin embargo, en los años setenta-ochenta, el modelo de Estado de bienestar entró en crisis debido a la insuficiencia de ingresos para solventar los gastos sociales, fue cuestionado y se propusieron nuevas combinaciones de lo público y lo privado en la provisión de servicios sociales (Gordon, 2000). Los gobiernos de AL optaron por el paradigma emergente, inspirado en el NPM, e introdujeron las llamadas reformas de “primera generación” o de ajuste estructural: desmantelamiento del intervencionismo estatal, aplicación de políticas de descentralización y de privatización de empresas públicas, reducción del tamaño del Estado y empleo del mecanismo de mercado como nueva centralidad. El Estado debía limitarse a garantizar las condiciones para que las empresas privadas ofrecieran los bienes y servicios públicos que habían constituido un componente fundamental de sus funciones en el modelo anterior, priorizando el objetivo de eficiencia, el establecimiento de prioridades claras, la reducción de costos administrativos, el incremento de la productividad de la inversión, etcétera (Gordon, 1999).

Al finalizar los ochenta, el anterior énfasis en políticas universales había sido sustituido por programas focalizados a la atención de la pobreza extrema, cuya medición en México (si bien difiere en función del método empleado) mostraba magnitudes alarmantes (véase cuadro 1).

En dicho marco, la nueva política social puede asumir objetivos de corto plazo, vinculados a la transferencia de recursos para paliar la situación actual de los pobres en relación con la satisfacción de sus necesidades elementales más inmediatas; o puede también plantearse objetivos de largo plazo, a través de la inversión en capital humano y social (educación, salud, alimentación, construcción de redes, fomento de los emprendimientos sociales, etcétera), en cuyo caso se habla de políticas de desarrollo social.

La aplicación de las reformas en la orientación del gasto y la inversión social llevó a que, en 2001, el Instituto Nacional de Estadística, Geografía e Informática (INEGI) reportara que la pobreza alcanzó a 57 millones de personas y la pobreza extrema a más de 10 millones (10.72 por ciento), mismas que podrían aumentar a 75 y 54 millones respectivamente, según los cálculos realizados por Boltvinik y Hernández (1999). En 2002, esto se combina con

CUADRO I  
 EVOLUCIÓN DE LA POBREZA EN MÉXICO  
 SEGÚN DIFERENTES AUTORES  
 (En porcentajes)

Año	<i>Hernández 1991</i>	<i>Boltvinik 1998</i>	<i>Boltvinik y Hernández 1999</i>	<i>INEGI 2001</i>	<i>Banco Mundial 2004a</i>
1963	77				
1968	72				
1977	58				
1984	59				
1994		69			
1996		78			
2001			75	59	
2002					52

Fuente: Elaboración propia con base en los datos de los autores citados.

uno de los índices de Gini más altos de América Latina: 0.531 (Naciones Unidas, 2002).

El Banco Mundial (2004b) informa un retroceso importante en el combate a la pobreza, provocado por la crisis de 1994-1995, que incidió para que se incrementara la pobreza general y extrema, llegando la última a 20.3 por ciento en 2002, a pesar del descenso de esta magnitud logrado en 1996-2002, con el que apenas se ha rebasado el nivel prevaleciente antes de la crisis mencionada. Michel Walton, Asesor Regional para la Reducción de la Pobreza y el Desarrollo Humano en el citado banco, señala como algunos de los factores que contribuyeron a la disminución registrada al "...crecimiento de las remesas internacionales para los pobres extremos en las zonas rurales y el aumento de la transferencias gubernamentales dirigidas a los pobres extremos a través de Oportunidades" (Banco Mundial, 2004b: 1).

Ante estos resultados, resulta necesario reorientar las políticas y programas sociales, rescatando los que han producido efectos positivos y desechando los que no lo han logrado, y valorando nuevas formas de gestión pública y social. Diversos autores (Kliksberg, 1997, 2000a y 2003, Nirenberg, 2003, Cunill, 1997, Bresser Pereira *et al.*, 2004, etcétera) destacan la importancia de generar procesos de participación social para mejorar la efectividad y ética de estos programas sociales; sin embargo, ésta presenta limitados avances.

Pero, ¿cuáles han sido los impactos de las distintas políticas y programas aplicados?, ¿cuáles son los problemas que les han impedido lograr sus

objetivos?, ¿qué cambios en su diseño e implementación permitirían mejorar su efectividad?, ¿qué otras alternativas podrían resultar más exitosas? Las respuestas a estas y otras preguntas similares requieren de la evaluación y aprendizaje a partir de la experiencia de los programas implementados, procesos que aún presentan limitados avances en México, como se verá más adelante.

Antes de pasar al análisis de dichas evaluaciones será necesario revisar brevemente los principales programas sociales que se aplicaron en el pasado y los que han estado vigentes durante el presente gobierno.

### Programas aplicados en sexenios previos

Desde el inicio de los setenta, el gobierno de Echeverría abandonó el modelo denominado de “desarrollo estabilizador” y lo sustituyó por el de “desarrollo compartido”, en que, como su nombre lo indica, se perseguía una mejor distribución de los frutos del desarrollo, a través de un incremento en la inversión pública dirigida a salud, educación, vivienda y, especialmente, al campo.

En este último se puso en marcha el Programa de Inversiones Públicas para el Desarrollo Rural (Pider), que puede ser considerado un antecedente de los programas de combate a la pobreza. Su objetivo consistía en concentrar la atención en las zonas más deprimidas, tratando de dinamizar el campo y disminuir la desigualdad.

López Portillo fue el primer presidente que aplicó programas sociales específicos. Continuó aplicando el Pider y, en 1977, propuso la creación de la Coordinación del Plan General de Zonas Deprimidas y Grupos Marginados (Coplamar), con la responsabilidad de: 1. realizar estudios para la atención de las necesidades esenciales de la población en las zonas deprimidas y grupos marginados, en el entendido de que su satisfacción constituía el verdadero sentido del desarrollo económico; y 2. efectuar propuestas para coordinar y mejorar el funcionamiento de las dependencias y entidades encargadas del desarrollo de los programas aprobados.

En este contexto se desarrollaron programas para la atención de cuatro millones de habitantes, como:

- Conasupo-Coplamar, dirigida a la población en general en materia de abasto de alimentos;
- Secretaría de Asentamientos Humanos y Obras Públicas (SAHOP)-Coplamar, para la provisión de agua potable y caminos rurales;

- IMSS-Coplamar, en la disposición de servicios médicos rurales;
- SEP-Coplamar, en la oferta de casas-escuela para niños de zonas marginadas alejadas;
- Secretaría de Trabajo y Previsión Social y Secretaría de Agricultura y Recursos Hidráulicos, STPS-SARH-Coplamar, vinculadas a la capacitación y el empleo para el fomento de los recursos naturales.

Coplamar enfrentó múltiples problemas, como la falta de recursos propios durante sus primeros tres años, mismos que podrían haber sido subsanados a tiempo de haber contado con evaluaciones sistemáticas. Al finalizar el sexenio, y a pesar de que se obtuvieron algunos resultados positivos, sólo siguieron aplicándose algunos programas sectoriales.

En el mismo sexenio también se creó el Sistema Alimentario Mexicano (SAM), para enfrentar la crisis que vivía el país en esa materia y con el objetivo de promover la autosuficiencia en alimentos básicos a través de créditos y subsidios a semillas y fertilizantes.

Durante el sexenio de De la Madrid se iniciaron los procesos descentralizadores y se puso en marcha el Programa de Nutrición y Salud, de corta duración (1987-1988), con el objetivo de mejorar el estado de las familias marginadas en las áreas señaladas.

El programa también enfrentó problemas de falta de coordinación entre las diversas instancias participantes a nivel federal, estatal y municipal: Secretaría de Salud (SSA), IMSS-Coplamar, INI, Instituto Nacional de la Nutrición, Conasupo, Diconsa, etcétera. Además contó con insuficiente tiempo para lograr algún impacto en materia de nutrición.

Desde los setenta había venido aumentando la crisis fiscal del Estado, lo que llevó al salinismo a adoptar más intensamente el enfoque que denominaría de liberalismo social, mismo que había empezado ya a aplicarse tenuemente durante el sexenio previo. Entre las medidas adoptadas sobresale la privatización de empresas estatales, que permitió contar con recursos para rediseñar la política social.

Es así que, en diciembre de 1988, se crea Programa Nacional de Solidaridad (Pronasol), con los objetivos de: 1. mejorar las condiciones de vida de los grupos campesinos, indígenas y colonos populares; 2. crear condiciones para el mejoramiento productivo de los niveles de vida de la población y promover el desarrollo regional; y 3. promover y fortalecer la participación y la gestión de las organizaciones sociales y de las autoridades locales. Se trata del primer programa focalizado exclusivamente hacia los más pobres, que impulsa la organización de las comunidades para la superación de sus problemas.

El mismo cubrió cuatro ámbitos específicos:

- solidaridad para el bienestar social, enfocado a los problemas de alimentación;
- solidaridad para la producción, para incrementar los ingresos de los pobres;
- solidaridad para el desarrollo regional, dirigida al desarrollo de infraestructura; y
- solidaridad en los programas sociales, en apoyo a la salud y la educación.

De esta manera, el programa pretendió aplicar recursos para aliviar situaciones graves de pobreza (sus efectos), pero al mismo tiempo atacar las causas de la misma a través del componente productivo y la disposición de infraestructura.

Se trató de un programa que incrementó fuertemente el gasto en desarrollo social: del 16.3 y 11.3 por ciento del gasto total federal ejercido en los sexenios previos al 24.8 por ciento. Su funcionamiento fue muy criticado por sesgos partidistas (generación de relaciones clientelares, asignación discrecional de recursos con fines electorales, etcétera), por sus altos costos y sus resultados poco contundentes. Sin embargo, también existen trabajos relevantes, como el de Székely (1993), en que se reivindica el esfuerzo financiero del gobierno, el componente participativo comunitario y se presentan ejemplos de experiencias exitosas.

Zedillo le dio continuidad a la política social de Salinas, a través del Programa de Educación, Salud y Alimentación (Progresá), enfatizando la importancia de la educación, ya que, en 1995, se reconocía la existencia de seis millones de personas analfabetas mayores de 15 años. El programa inició actividades en agosto de 1997 y se dirigió a los pobres extremos (con ingresos menores a un salario mínimo mensual de 1,003 pesos), que habitaban zonas urbanas de alta marginación. Combinó incentivos y también sanciones por incumplimiento de obligaciones, y ofreció tres tipos de apoyos: becas para los niños entre tercero de primaria y tercero de secundaria, apoyos monetarios para el consumo familiar y consultas gratuitas de salud. Contó con menos recursos que su antecesor y fue incapaz de ofrecer todos los empleos bien remunerados que reconoció se necesitaban para la superación de la pobreza, por lo que debió proceder a la aplicación del programa especial de Empleo Temporal.

Progresá trató de superar las críticas que se habían dirigido a su antecesor estableciendo criterios explícitos de selección de beneficiarios con base en estudios sociodemográficos y un sistema de puntajes en variables relativas a ingreso,

educación, vivienda, etcétera. En cambio, no explicitó modalidades de coordinación con las instituciones educativas y de salud para ofrecer servicios de mejor calidad, y menos aún en materia de vivienda, agua potable o capacidades para el trabajo (Gordon, 1999).

En síntesis, los ejes fundamentales que guiaron la política social del periodo fueron: 1. focalización; 2. descentralización (aunque con financiamiento y criterios centralizados); y 3. pequeña participación de las autoridades locales y los destinatarios para apuntalar su legitimidad. Los programas para la superación de la pobreza del periodo complementaron la inversión en el desarrollo de capacidades básicas de Progresá con: 1. inversión en infraestructura básica en el entorno (agua potable, electricidad, etcétera) y en condiciones que favorecieran la producción de nuevas oportunidades de ingreso (apoyo a proyectos productivos); y 2. programas de empleo dirigidos a su generación y a la capacitación para el trabajo (Gordon, 1999).

A pesar de la aplicación de los programas revisados, la distribución del ingreso entre grupos sociales y regiones geográficas siguió polarizándose e intensificándose la pobreza urbana, por lo que a continuación se revisa en forma más amplia la política social del actual gobierno de Fox.

## Plan Nacional de Desarrollo 2001-2006

El Plan Nacional de Desarrollo (PND) constituye el documento rector fundamental que guía la acción gubernamental durante el sexenio. Por esa razón se aludirá al mismo con objeto de identificar los principales problemas, objetivos y estrategias planteados en relación al mejoramiento del bienestar social de los ciudadanos, la importancia atribuida a la evaluación de las políticas y programas sociales y el papel que se le otorga a la participación social.

Sus objetivos incluyen: consolidar el avance democrático, reducir la inseguridad y cancelar la impunidad, abatir la pobreza y lograr una mayor igualdad social, reformar la educación para asegurar oportunidades de educación integral y de calidad para todos los mexicanos, alcanzar un crecimiento económico con estabilidad, incluyente y justo, asegurar la transparencia y rendición de cuentas del gobierno y descentralizar las facultades y recursos de la Federación. Considera como palancas del cambio a lograr: la educación, el empleo, la democratización de la economía, el federalismo y el desarrollo regional (PND, 2001: 12).

Mejorar el bienestar social de los mexicanos se considera el fin último perseguido por el PND. Lograrlo implica superar amplios, profundos y añejos problemas, de los que refiere la siguiente información, correspondiente al año 2000:



- pobreza: 55 por ciento de la población recibía un ingreso inferior a dos salarios mínimos mensuales;<sup>38</sup>
- desigualdad económica: el 20 por ciento de la población de mayores ingresos concentraba el 58 por ciento de los mismos, mientras que el 20 por ciento más pobre recibía el 3.7 por ciento del ingreso total del país;
- vivienda: 47.5 por ciento vivían en condiciones de hacinamiento, el 13.3 por ciento tenía pisos de tierra;
- educación: 27.8 por ciento de la población mayor de 15 años de edad no había terminado la primaria, el promedio de escolaridad de la población era de 7.56 grados, y sólo el 20 por ciento de los jóvenes tenía acceso al nivel superior;
- servicios: el 15.4 por ciento no contaba con agua entubada y el 21.5 por ciento carecía de drenaje.

Esta situación se agravaba en el caso de grupos vulnerables como los cerca de 10 millones de indígenas que habitan el país, cuya tasa de analfabetismo equivalía en 1995 a 3.6 veces el promedio nacional (10.6 por ciento), un 20 por ciento menos que el promedio nacional finalizaba la educación primaria, y su mortalidad infantil era de 48.3 por 1,000 niños nacidos, frente a los 29.0 de la media general del país. Algo similar sucedía con los 2.2 millones de personas que sufren algún tipo de discapacidad física o mental, y demás grupos vulnerables.

A todo lo anterior se agregaba la insatisfacción de las familias mexicanas con la calidad de los bienes y servicios ofrecidos por los programas gubernamentales de salud y educación, la inequidad en su acceso, la calidad de la infraestructura, etcétera.

Para enfrentar los problemas anteriores, el gobierno foxista creó la Comisión para el Desarrollo Social y Humano, cuyo objetivo central incluyó: “...coordinar las inversiones en justicia social, eliminar los desequilibrios sociales extremos y procurar una vida digna para todos, promoviendo la iniciativa individual y colectiva de los mexicanos, en especial para aquellos que, por tiempo inmemorial, aguardan la justicia y cotidianamente sufren la miseria, el abandono, la ignorancia y la violencia” (PND, 2001: viii).

La revolución educativa es la estrategia prioritaria del gobierno, complementada por un sistema integral de salud; uno de sus postulados fundamentales es la equidad, que se aparece confundido con la igualdad de oportuni-

<sup>38</sup>El salario mínimo diario correspondía en el año 2000 a 37.90, 35.10 y 32.70 pesos en las zonas geográficas A (incluye al Distrito Federal), B y C, respectivamente.

dades. Las acciones en materia de política social giran alrededor de seis ejes, en los que se establecen los aspectos a evaluar:

1. Evitar la existencia de grupos de población “...cuyas condiciones de vida, oportunidades de superación personal y de participación social, se encuentren por debajo de ciertos umbrales” (PND, 2001: 50). Los aspectos a medir, en cantidad y calidad, serían: educación, salud, vivienda, infraestructura y reducción de la pobreza.
2. Ofrecer equidad en los programas e igualdad de oportunidades para permitir la superación del nivel de vida de las personas y grupos más vulnerables y numerosos (indígenas, niños, ancianos, discapacitados, mujeres, jóvenes), sin abandonar los programas de cobertura general. La evaluación de este eje se realizaría con base en información sobre las oportunidades en el acceso a la salud, a la educación, al trabajo y en otros aspectos del ámbito social y humano que tengan distintos grupos de la población, incluyendo las diferencias por género.
3. Fomentar la actitud independiente y emprendedora de los ciudadanos, su capacidad e iniciativa, promoviendo proyectos que mejoren la preparación, la escolaridad y los conocimientos de la población, conduzcan al desarrollo de sus habilidades y destrezas, etcétera. En este eje se evaluaría la cobertura, calidad y repercusión de los programas, mediante información estadística relativa a escolaridad, competencias laborales y sociales, y capacidades emprendedora y de innovación.
4. Propiciar, conservar y alentar la cohesión social, aumentar la solidaridad de los mexicanos entre sí y con el bien común, etcétera, mediante políticas y proyectos incluyentes. En particular, fomentar la participación de las organizaciones sociales y civiles en el diseño, instrumentación y evaluación de las políticas y programas públicos. La evaluación se basaría en información sobre procesos de participación comunitaria, funcionamiento de organizaciones civiles y sociales, y existencia de valores solidarios.
5. Construir una cultura del desarrollo en armonía con la naturaleza, que fomente el conocimiento sobre el deterioro del medio ambiente y su impacto en el bienestar. Su evaluación se realizará a partir de información sobre la moderación del daño a la atmósfera, el consumo de energía, la pérdida de sistemas forestales, la tasa de conservación de acuíferos, etcétera.
6. Recuperar la confianza en la capacidad del gobierno y las instituciones, mediante, entre otras acciones, el diseño de respuestas eficaces a

situaciones no previstas de orden natural y social. Para su evaluación se previó la utilización de indicadores como la rendición de cuentas, la calidad del trato a los usuarios de los servicios públicos y el alcance de la protección financiera frente a eventos catastróficos (PND, 2001: 48-53).

En realidad el documento se refiere a los “indicadores” a utilizar en la evaluación de cada uno de estos ejes, pero se limita a señalar variables muy generales y abstractas sobre las que se recabará información, ya que serán las organizaciones mismas las encargadas de diseñar sus indicadores específicos.

Uno de los cinco objetivos de las comisiones creadas por el gobierno se relacionaba con la *evaluación de las políticas y programas sociales*:

Establecer procesos de evaluación de las acciones de gobierno que ajusten su labor a las necesidades de la sociedad” (p. 66). ... “La transparencia en el uso de los recursos y la ponderación objetiva de las consecuencias de las políticas públicas constituyen prácticas que deberán generalizarse en México. En materia social, además de alcanzar metas cuantitativas cada vez más ambiciosas, es necesario lograr incrementos de calidad y eficiencia que redunden en mayor satisfacción de la población y en mayor productividad de los recursos presupuestales destinados a su atención (p. 94).

Ante todo, se prevé que las políticas y los programas sociales sean evaluados por la propia administración pública federal, quien integrará sus avances, logros y retrasos en un informe de evaluación del PND a preparar anualmente (p. 9), de igual forma que habrían participado en el proceso de planeación.

El gobierno declaró su vocación de dialogar en materia de trabajo, desarrollo social, grupos vulnerables y competitividad, tanto con el Congreso de la Unión y el Poder Judicial, como con las autoridades de otros ámbitos de gobierno. Sus evaluaciones deberían servir como contrapeso del Poder Ejecutivo y se complementarían con la participación de la ciudadanía en estos procesos, misma que será considerada con mayor detalle en el siguiente apartado.

El Sistema Nacional de Planeación Participativa (SNPP) se compone de tres procesos: planeación estratégica, seguimiento y control, y mejoramiento organizacional. Si bien aquí ya no hay una clara referencia a la evaluación, la misma se vincula al segundo proceso señalado, que prevé el desarrollo de un subsistema de análisis, seguimiento y control, cuyo elemento central es

el Sistema Nacional de Indicadores, que cada secretaría y organismo descentralizado diseñaría para verificar que se cumpliera con las demandas de la sociedad y los objetivos y metas del PND, aplicando medidas cuando se perciban desviaciones.

Para ello se propuso construir una cultura de la información sobre los resultados de la administración pública, sus oportunidades y retos que acompañada de la claridad en las metas y objetivos de los programas, permitiera su seguimiento y evaluación. La misma debería complementarse con procesos de capacitación y superación continua de los servidores públicos.

En particular, el PND destaca la importancia atribuida a la participación ciudadana en la conciliación de intereses individuales y sociales, de opiniones de especialistas y otros ciudadanos, de la experiencia y el sentido innovador. Entre otras el SNPP:

...impulsará un proceso de definición, concertación, seguimiento y evaluación de las políticas y acciones del Poder Ejecutivo Federal y las actividades de todas las dependencias y entidades de la administración pública federal, además de integrar la opinión de la población, mediante mecanismos de participación ciudadana para la elaboración y evaluación de planes y programas (PND, 2001: 9)... El PND establece como condición fundamental un diálogo abierto y permanente con los ciudadanos, con las agrupaciones políticas y sociales, con los otros poderes y ámbitos de gobierno, y con la sociedad en general (*Ibidem*: 13).

Este proceso de participación se consideró iniciado durante los meses posteriores a las elecciones del 2 de julio de 2000, en que ciudadanos y grupos de la sociedad se habrían acercado al Presidente electo para expresar sus opiniones y propuestas, y diferentes coordinaciones habrían realizado múltiples reuniones de consulta diagnóstica y propositiva.

Se pretendió continuar con el mismo a partir de diciembre de 2000, en que el licenciado Fox asumió la Presidencia, mediante la aplicación de encuestas y reuniones de opinión ciudadana para recoger su contribución a la elaboración del PND. La utilización de estos dos procesos permitió una participación total de 174,865 personas (menos del 2 por ciento de la población del país) que presentaron 379,525 propuestas (PND, 2001: 15).

El PND reconoce también importantes cambios en la participación social: mientras hasta la finalización de la década de los sesenta, las organizaciones de la sociedad civil eran sólo de carácter sindical o sectorial, a partir de esa

fecha se ha multiplicado el número de organizaciones de la sociedad civil interesadas en los asuntos públicos.

Hasta aquí la descripción que realiza el PND respecto de las primeras etapas del gobierno foxista, que dan una imagen de alta participación ciudadana, muy discutible. El mismo documento se compromete a pasar de la “amplia participación ciudadana en la elaboración del PND a una amplia participación en los procesos de evaluación de resultados...” (PND, 2001: 45). Resta por analizar lo que pasó posteriormente en relación con los mecanismos previstos para dicha evaluación participativa de las políticas, programas y acciones gubernamentales.

### Programa Nacional de Desarrollo Social 2001-2006

El Programa Nacional de Desarrollo Social (PNDS) establece los objetivos, ejes rectores, estrategias y líneas de acción a desarrollar en materia de desarrollo social. Se enfoca a la superación de la pobreza externa como tarea inmediata e impostergable, ya que reconoce que el crecimiento económico no se ha traducido en bienestar para toda la población. Sin embargo, al definir la magnitud de la pobreza, plantea que es prácticamente imposible que el Estado pueda atenderla solo, por lo que requiere de la participación de otros sectores sociales (PNDS: 21).

Pretende dotar a la población de nuevas capacidades mediante un enfoque integral, compensatorio, focalizado y multidimensional, rechazando el asistencialismo en favor de la subsidiariedad y la participación de los beneficiarios en la toma de decisiones. Plantea que la política social no debe estar subordinada a la económica y debe adquirir la característica de política de Estado, con una perspectiva de largo plazo.

Por otro lado, señala el avance de los procesos de globalización y competitividad, para la cual el país necesita resolver los rezagos educativos y las condiciones de pobreza, por lo que la estrategia de Progreso se considera adecuada y se mantiene como Programa de Desarrollo Humano Oportunidades. En cambio se hace una crítica de la política social previa por estar desvinculada de la política económica.

Para lograr los objetivos y metas propuestas, la Secretaría de Desarrollo Social (Sedesol) se compromete a realizar una evaluación integral de los programas sociales que permita conocer el desempeño de los mismos, recomendar alternativas para mejorar y promover la rendición de cuentas, incluyendo seis tipos de evaluación: 1. de diagnóstico; 2. de seguimiento; 3. de gestión, 4. participativa; 5. de resultados; y 6. de impactos.

También se retoma la propuesta del PND de creación de un sistema de indicadores de seguimiento de los programas y se plantea aplicar un esquema de mejora continua de los mismos, que permita dar respuesta inmediata a los cambios que se presenten en la demanda social y a las limitaciones de la oferta institucional (PNDS: 160).

### Estrategias Contigo y Microrregiones

La *estrategia Contigo* constituye la base de la política social del sexenio y pretende transformarse en política de Estado. Se propone impulsar y lograr el desarrollo humano integral de los ciudadanos y responder a los graves rezagos que enfrentan amplios sectores de la población. Declara promover la integración y articulación de las acciones para potenciar el uso de los recursos y maximizar su impacto (SHCP, 2003). Se fundamenta en un conjunto de postulados: buen gobierno, búsqueda de impactos, manejo transparente, etcétera.

Sus objetivos son cuatro:

- reducir la pobreza extrema;
- apoyar el desarrollo de las capacidades de las personas en condición de pobreza;
- generar igualdad de oportunidades para los grupos más pobres y vulnerables; y
- fortalecer el tejido social, fomentando la participación y el desarrollo comunitario.

En términos generales la estrategia busca:

- Articular los programas federales entre sí y con los que desarrollan los gobiernos estatales, municipales, las organizaciones civiles y las comunidades;
- crear nuevos programas en áreas que, hasta ahora, habían sido consideradas de baja prioridad, pero que son necesarias a fin de tratar de manera integral la superación de la pobreza;
- transformar acciones que requieren un renovado impulso con objeto de aumentar su impacto bajo el nuevo esquema; y
- reforzar los programas existentes con potencial para aportar sus beneficios, con apego a sus principios, en las cuatro vertientes siguientes:
  - desarrollo de capacidades,
  - generación de oportunidades,

- desarrollo del patrimonio, y
- protección social.

Las anteriores se vinculan con las tres definiciones de pobreza manejadas por la actual política social:

- Pobreza alimentaria: se considera que las personas y los hogares viven en esta condición cuando, aun si asignan todo el ingreso que perciben para cubrir sus necesidades de alimentación, éste es insuficiente para garantizar el consumo mínimo establecido en una canasta alimentaria normativa. Estas familias, por tanto, no cuentan siquiera con los recursos mínimos para la adquisición de los alimentos necesarios para su subsistencia, tampoco para desarrollar sus capacidades o cubrir cualquier necesidad adicional como vestido o vivienda.
- Pobreza de capacidades: se refiere a las capacidades humanas mínimas que permiten llevar una vida saludable y productiva y se considera que las personas y los hogares viven en condiciones de pobreza de capacidades cuando sus ingresos son insuficientes para cubrir conjuntamente sus necesidades básicas de alimentación, salud y educación.
- Pobreza de patrimonio: la que enfrentan las personas y los hogares cuando sus ingresos no son suficientes para satisfacer, en conjunto, sus necesidades de alimentación, salud, educación, vivienda, vestido y transporte.<sup>39</sup>

Este diseño que reconoce múltiples dimensiones de la pobreza y de la acción pública consecuente y toma en cuenta el ciclo de vida de los individuos al interior de sus hogares, le ha valido el calificativo de excelente por parte del Banco Mundial (1994a).

Para cumplir con la estrategia prioritaria del PND de lograr el desarrollo social y económico de los sectores más marginados, la Sedesol definió la estrategia Microrregiones, y comenzó por mapear a nivel municipal y local la ubicación de la marginación en México. Se identificaron 263 regiones prioritarias, integradas por 99,891 localidades (62,238 de ellas con menos de 50 habitantes), pertenecientes a 1,334 municipios con alta y muy alta marginación y énfasis en población hablante de lenguas indígenas, en las 31 entidades federativas, con una población estimada en

<sup>39</sup>Se trata de las necesidades fisiológicas y de seguridad del ser humano, que según la pirámide de necesidades elaborada por Abraham Maslow (1954), *Motivation and personality*, Nueva York, Harper & Row Publishers Inc., incluye también las de afiliación o aceptación social, de autoestima y de autorrealización.

19'225,833 habitantes (26 por ciento hablante de lenguas indígenas, en condiciones de pobreza). Se aclara que sólo cinco estados del país carecían de municipios en condiciones de marginación y habitados por población indígena: Aguascalientes, Baja California, Baja California Sur, Coahuila y Colima. Para ellos se definió un techo diferente de marginación y se incluyeron 13 de sus municipios, con lo cual la estrategia cubre todos los estados del país.

La estrategia Microrregiones implica la unión de esfuerzos de los tres órdenes de gobierno y la sociedad civil para abatir los rezagos en los territorios que presentan marginación crónica, con la participación activa de las comunidades en su propio desarrollo. A nivel federal cuenta con la participación de 14 secretarías, encargadas de la operación de programas concretos, lo que requiere de un importante esfuerzo de coordinación y una gran capacidad de argumentación y convencimiento para romper con la sectorialidad largamente arraigada.

Además, Microrregiones se presenta como una respuesta a la búsqueda de una metodología alternativa para balancear el trabajo que se hacía al inicio de la presente administración, de corte economicista y basado en políticas subsidiarias y personalizadas (tipo Pronasol, dirigido a grandes núcleos de población para la obtención de mayor rédito electoral).

En este marco, la participación ciudadana se considera responsable de validar lo que se va haciendo, ya que las iniciativas deben responder a las problemáticas y objetivos locales; en el largo plazo se espera que los ciudadanos sean capaces de apropiarse de la estrategia.

El enfoque territorial busca generar impactos en la infraestructura y servicios de Microrregiones, en consecuencia, requiere de una evaluación especialmente diseñada a tal efecto.

Los estados fueron los encargados de proponer formas de aglutinar los municipios en las microrregiones. Se identificaron lugares estratégicos en cada microrregión, seleccionados con base en criterios de volumen y estabilidad de la población, conexión con su entorno, disponibilidad de servicios educativos, de salud, infraestructura básica, vocación productiva, etcétera, con la intención de ubicar centros estratégicos comunitarios (CEC) para que la localidad se transforme en un "polo de desarrollo" para su entorno y vaya alcanzando las llamadas 11 "banderas blancas".

Comenzó a operar en 2002, cubriendo 2,966 localidades, en 539 municipios de muy alta marginación, con más de un 40 por ciento de la población hablante de lenguas indígenas, ubicados en 18 estados y 155 microrregiones.



CUADRO 2  
PRESUPUESTO EJERCIDO EN DESARROLLO SOCIAL  
POR VERTIENTE, 2001-2002

<i>Vertiente</i>	<i>Presupuesto 2001 (millones de pesos)</i>	<i>Presupuesto 2002 (millones de pesos)</i>
Desarrollo de capacidades	29,620.3	35,743.2
Desarrollo del patrimonio	22,213.8	28,046.1
Generación de oportunidades	10,683.7	13,900.1
Protección social	186.8	388.3
Totales	62,704.6	78,077.7

Fuente: SHCP, Cuenta de la Hacienda Pública 2002.

En los anteriores se aplicaron 68 programas sociales.<sup>40</sup> El gasto ejercido por los programas ha sido, según el PEF, 2003, el más alto de la historia del país, tanto como proporción del gasto programable como en términos reales.<sup>41</sup> Los montos correspondientes a 2001 y 2002 se presentan en el cuadro 2.

A continuación se desglosa la presentación en las tres vertientes en que participó Sedesol en 2001-2002, ya que la última es desarrollada exclusivamente por la SSA y el IMSS, y se mencionan algunas modificaciones introducidas en 2003.

### *Desarrollo de capacidades*

Esta vertiente incluye al programa Oportunidades y otros programas alimentarios, de educación y de salud. Sedesol participa en el primero y, en 2002, se ocupó de los programas de Subsidio a la Tortilla, Abasto Social de Leche y Abasto Rural, dentro de los alimentarios.

Los presupuestos ejercidos por Sedesol se presentan en el cuadro 3. La disminución del gasto de Liconsa se explica por la depuración continua que sufrió el padrón de beneficiarios de Tortilla a lo largo del año.

El programa de Desarrollo Humano Oportunidades constituye la continuación del Progreso del sexenio anterior, con algunas acciones nuevas y otra fortalecidas, que busca ofrecer mayores opciones para que las familias más

<sup>40</sup> El Banco Mundial (junio de 2004), menciona 104 programas que ofrecen subsidios públicos.

<sup>41</sup> Según el Banco Mundial (2004), entre 2000 y 2002, el gasto en desarrollo social se incrementó en promedio 5.3 por ciento al año y el dirigido especialmente a reducción de la pobreza en 14.2 por ciento, frente a sólo un 6.3 por ciento que correspondió al gasto total. Dicho 14.2 por ciento representa el 1.3 por ciento del PIB, en comparación con el 0.7 por ciento que se dedicaba en 1990.

CUADRO 3

PRESUPUESTO EJERCIDO POR PROGRAMAS DE LA VERTIENTE  
DESARROLLO DE CAPACIDADES, 2001-2002

<i>Programas</i>	<i>Presupuesto 2001 (millones de pesos)</i>	<i>Presupuesto 2002 (millones de pesos)</i>
Oportunidades (sólo componente de Sedesol)	5,105.7	6,864.5
Tortilla y Abasto Social de Leche (Liconsa)	1,389.7	774.5
Abasto Rural (Diconsa)	510.7	568.0
Totales	7,006.1	8,207.0

Fuente: Elaboración propia a partir de: SHCP, Cuenta de la Hacienda Pública 2002. Sedesol informó por vía del IFAI que el gasto de Oportunidades había sido de 5,965.60 pesos, y el de Abasto Rural de 6,699.50 pesos, significativamente diferentes de los aquí presentados por la SHCP.

pobres mejoren sus niveles de bienestar, mediante acciones intersectoriales coordinadas de salud, educación y desarrollo social.

El programa cuenta con una coordinación nacional dependiente de la Secretaría de Desarrollo Social y funciona como órgano desconcentrado. El titular de la coordinación es nombrado por el Presidente de la República. Cuenta con un consejo formado por los titulares de las secretarías involucradas en su operación (Sedesol, SSA y SEP) y la SHCP. Este consejo está apoyado por un comité técnico integrado por los subsecretarios de las mismas y de la Secretaría de la Función Pública, así como por representantes del IMSS y la Coordinación Nacional. Cuenta con cuatro componentes:

1. el componente alimenticio, que entrega subsidios económicos y papi-las para niños menores de cinco años y mujeres embarazadas y en periodo de lactancia;
2. el educativo, que otorga becas a niños y jóvenes inscritos en el ciclo escolar a nivel primaria, secundaria y educación media superior;
3. el de salud, que ofrece consultas médicas a la familias en control; y
4. el patrimonial, que conforma un fondo de ahorro para cuando el beca-rio concluya sus estudios medios superiores.

En 2002 funcionó en 70,520 localidades de 2,354 municipios y de las 263 regiones prioritarias de atención inmediata, incorporó a su padrón de beneficiarios más de un millón de nuevas familias alcanzando una cobertura de 4.2 millones, y recibió el 47.6 por ciento del total de recursos de

la vertiente, si se considera la participación de Sedesol, SSA y SEP en la misma.

El programa Abasto Social de Leche funciona desde 1949 y pasó a ser operado por Liconsa desde su creación en 1972. Hasta enero de 1995 formó parte del sector gubernamental de comercio y, en esa fecha, se integró al sector social, a cargo de la Sedesol. Entre 1990 y 1995 Liconsa aplicó una estrategia de desincorporación que redujo su número de plantas de 18 a nueve (buena parte de las plantas pasteurizadoras, las tres plantas industrializadoras y los centros de acopio), restringiendo su actividad productiva exclusivamente a leche reconstituida. Su abasto ha estado históricamente concentrado en el Distrito Federal y el Estado de México, ampliándose desde 1996 a otros estados del centro y occidente del país (Del Valle y Roldós, 2003).

Actualmente, Liconsa ofrece leche a precios subsidiados a las familias en condiciones de pobreza extrema, que no cuenten con apoyo del programa Oportunidades, con niños menores de 12 años, que habitan en zonas urbanas, semiurbanas y rurales marginadas; así como a adultos mayores de 60 años, enfermos y/o discapacitados mayores de 12 años y a mujeres en periodo de gestación, en condiciones severa de marginación. También hace entrega de su producto, mediante convenios, a instituciones como el INI y al programa de Atención a Jornaleros Agrícolas, entre otros.

Su cobertura en 2002 fue de 4.8 millones de beneficiarios en promedio diario; 27.8 por ciento de ella, ubicada en localidades rurales.

El programa de Subsidio a la Tortilla fue creado en 1986 y estuvo a cargo del Fideicomiso para la Liquidación al Subsidio de la Tortilla (Fidelist) hasta noviembre del 2000, en que pasó a ser responsabilidad de Liconsa.

En 2002, ofreció dotaciones de un kilogramo diario y gratuito de tortillas a las familias en condiciones de pobreza extrema que habitan en colonias marginadas y no son beneficiarias de otros programas federales, ubicadas en localidades de más de 15,000 habitantes.

Durante 2002, el padrón registró una disminución del 39.7 por ciento, alcanzando una cobertura de 633,172 familias, proceso que finalizaría con el cierre del programa el 31 de diciembre de 2003.

Por último, Diconsa distribuyó alimentos dentro del programa Abasto Rural, a través de 22,516 tiendas (1.5 por ciento menos que el año anterior), ubicadas en 1,334 municipios. De las anteriores, 12,802 se encuentran en las microrregiones de atención inmediata.

La cobertura informada para 2001 y 2002 fue de 29,152 familias.

### *Desarrollo del patrimonio*

Su objetivo es que las familias tengan certeza en la propiedad jurídica de sus activos y dispongan de una vivienda digna con una infraestructura suficiente de agua potable, drenaje, alcantarillado y caminos, entre otros.

En 2002, Sedesol participó, entre otras secretarías, encargándose directamente del desarrollo de los programas del Ramo 20 (su ubicación en esta vertiente no siempre resulta clara) cuyos presupuestos se presentan en el cuadro 4.

El programa de Ahorro y Subsidios para la Vivienda Progresiva dio seguimiento a las actividades largamente desarrolladas en el marco del Fondo Nacional para las Habitaciones Populares (Fonhapo). Se destinó a la dotación de vivienda y a la ampliación y mejoramiento de la infraestructura social básica de las zonas y microrregiones de alta y muy alta marginación. En 2002, abarcó 24 entidades federativas en materia de adquisición de vivienda y 22 en cuanto a su mejoramiento.

Su población objetivo está constituida por las familias de la población urbana con ingresos de hasta tres veces el salario mínimo mensual, que no han sido atendidos por los organismos nacionales de vivienda y no tienen acceso al crédito de la banca comercial.

Desarrollo de los Pueblos y Comunidades Indígenas es un programa que se inició en 2002, en 20 entidades federativas, con el objetivo de provocar el impulso social y económico de los pueblos y comunidades indígenas, mediante la generación de proyectos de infraestructura de carácter regional

CUADRO 4

#### PRESUPUESTO EJERCIDO EN PROGRAMAS DE LA VERTIENTE DESARROLLO DEL PATRIMONIO, 2001-2002

<i>Programas</i>	<i>Presupuesto 2001 (millones de pesos)</i>	<i>Presupuesto 2002 (millones de pesos)</i>
Ahorro y Subsidios para la Vivienda Progresiva	437.9	897.9
Desarrollo de los Pueblos y Comunidades Indígenas	0.0	824.7
Estatales por Demanda	0.0	368.2
Microrregiones	0.0	220.0
Zonas de Alta Marginación e Indígenas	67.5	101.6
Totales	505.4	2,412.4

Fuente: SHCP, Cuenta de la Hacienda Pública 2002. Sedesol informó por vía del IFAI que el gasto del primer programa reportado había sido de 882.47 pesos, mínimamente diferente del aquí presentado por la SHCP.

y otros de alcance productivo, que respondan a sus demandas e impacten significativamente en el mejoramiento de sus condiciones de vida, con pleno respeto a los recursos naturales de su entorno, a su cultura y derechos.

El programa Estatales por Demanda también dio inicio en 2002, para sustentar la ejecución de proyectos estatales que hayan demostrado su viabilidad técnica y social, en concordancia con la estrategia de superación de la pobreza instrumentada por las 30 entidades federativas en que se aplicó, de acuerdo con las prioridades establecidas en los planes estatales de desarrollo.

Microrregiones constituye un programa específico (denominado Desarrollo Local a partir de 2003), que forma parte de la estrategia del mismo nombre, tiene presencia en todos los estados del país y promueve el desarrollo integral y sustentable de los espacios territoriales que registran los índices más altos de marginación del país, adicionando aquellos municipios en los que la población indígena sea predominante, a través de la ejecución de proyectos estratégicos emanados directamente de un proceso de planeación participativa, que fomente de manera corresponsable el desarrollo económico, social y humano de su población.

Zonas de Alta Marginación e Indígenas comenzó a funcionar en los estados de Chiapas, Oaxaca y Guerrero. Su objetivo era contribuir a la generación de procesos complementarios de desarrollo regional integral en las comunidades en pobreza, que muestren expresiones agudas de exclusión social y resquebrajamiento del tejido social mediante la realización de proyectos estratégicos sustentables, con la participación de los gobiernos estatales y municipales y la sociedad.

Iniciativa Ciudadana 3×1 fue creado en 2002. Si bien no aparece en el cuadro general en que se detalla el presupuesto ejercido, se agrega en su desglose refiriendo una inversión federal de 99.6 millones de pesos para 2002, aplicada en 21 entidades federativas.

El programa busca respaldar las iniciativas ciudadanas para concretar proyectos que conlleven a mejorar la calidad de vida de los habitantes mediante la concurrencia de recursos de los tres órdenes de gobierno y de los propios ciudadanos organizados, radicados en el país o en el extranjero.

### *Generación de oportunidades*

Esta vertiente se propone impulsar el desarrollo de proyectos productivos que coadyuven a generar oportunidades de empleo e ingreso a la población que habita en regiones de muy alta y alta marginación.

En 2002, Sedesol participó en la misma, haciéndose cargo directamente de los programas de Ramo 20 de Oportunidades Productivas y Atención a Jornaleros Agrícolas y, uniendo sus esfuerzos conjuntamente con la Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación (Sagarpa), la Secretaría de Medio Ambiente y Recursos Naturales (Semarnat) y la Secretaría de Comunicaciones y Transporte (SCT), en el programa de Empleo Temporal.

Los presupuestos ejercidos por Sedesol en estos programas se muestran en el cuadro 5.

Oportunidades Productivas (llamado Opciones Productivas a partir de 2003) es de reciente creación, se desarrolla en todos los estados del país y persigue la finalidad de fortalecer el patrimonio productivo bajo un esquema que garantiza la capitalización, mediante el impulso a la producción sustentable y la diversificación de actividades adecuadas a las condiciones de zonas de temporal con baja productividad y/o alta siniestralidad.<sup>42</sup>

Atención a Jornaleros Agrícolas se ha aplicado en 17 entidades federativas para mejorar las condiciones de vida y de trabajo de la población jornalera agrícola, proporcionando una atención integral a través de procesos de coordinación institucional con los tres órdenes de gobierno y de concertación social con productores, organismos sociales y los mismos beneficiarios.

Empleo Temporal ofrece cobertura en todos los estados del país con el propósito de generar oportunidades de empleo e ingresos en el medio rural a efectos de abatir el rezago económico y social, al utilizar la mano de obra no calificada en las zonas de mayor marginación, con acciones

CUADRO 5

PRESUPUESTO EJERCIDO POR LOS PROGRAMAS DE LA  
VERTIENTE GENERACIÓN DE OPORTUNIDADES, 2001-2002

<i>Programas</i>	<i>Presupuesto 2001 (millones de pesos)</i>	<i>Presupuesto 2002 (millones de pesos)</i>
Oportunidades productivas	527.4	910.1
Atención jornaleros agrícolas	135.8	173.0
Empleo temporal	ND	1,522.1
Totales	ND	2,605.2

Fuente: Elaboración propia a partir de SHCP, *Cuenta de la Hacienda Pública 2002*.

<sup>42</sup> Este programa fue calificado por el director general de evaluación como teniendo exclusivamente objetivos políticos (Hernández, 2005).

dirigidas al desarrollo del capital humano y de la infraestructura social y productiva, así como para la conservación y aprovechamiento de los recursos naturales.<sup>43</sup>

En 2002, se desarrolló en su modalidad tradicional, con 1,378.4 millones de pesos de gasto ejercido y en la modalidad de emergencia, con los restantes 143.7 millones. La SHCP informa que se crearon 418,792 empleos temporales en zonas rurales marginadas dentro de la primera modalidad y 75,506 en obras de reparación de daños causados por desastres naturales, en ocho estados, dentro de la segunda.

### Otros programas

Además de los programas mencionados en la estrategia Contigo, en 2002, Sedesol desarrolló otros programas, sectorizados o directamente aplicados en el marco del denominado Ramo 20.

- Los programas sectorizados fueron operados por institutos, comisiones o fideicomisos:
- Se dirigieron especialmente a población indígena y estuvieron a cargo del INI siete programas: Promoción de Culturas Indígenas; Desarrollo Económico Productivo de los Pueblos Indígenas; Planeación, Bienestar Social, Reparación y Equipamiento de Albergues; Operación de Albergues; y Promoción y Procuración de Justicia.
- Dependieron de Conaza, que atiende prioritariamente a las comunidades rurales de las zonas áridas y semiáridas del país que viven en pobreza extrema (especialmente en 17 entidades), dos programas y un plan: Programa de Vinculación y Transferencia Tecnológica, Plan Acuario, y Programa Para el Desarrollo de los Habitantes del Semidesierto.
- El Fonart fue responsable del programa de Compras, Créditos, Concursos y Asistencia Técnica, cuyo objetivo general es apoyar a los artesanos de México.
- El Instituto Nacional de Personas Adultas Mayores (Inapam) contribuyó con el programa Expertos en Acción, dirigido a toda persona jubilada y/o pensionada con la capacidad de participar en las acciones dirigidas a grupos de población que subsisten en condiciones de alta vulnerabilidad.

<sup>43</sup>Al igual que en el caso de Oportunidades Productivas (posteriormente Opciones Productivas), el director general de evaluación menciona que se trata de uno de los programas que aún tiene sólo objetivos políticos (Hernández, 2005).

–Finalmente, Indesol desarrolló cuatro programas: Coinversión Social, Investigación Para Desarrollo Local, Capacitación Para el Fortalecimiento de la Sociedad Civil y Fortalecimiento Institucional Para la Gestión Social Municipal.

Dentro del Ramo 20, la Sedesol, a través de la Dirección General de Políticas Sociales, gestionó otros cuatro programas: Identidad Jurídica, Jóvenes por México, Mujeres Jefas de Familia, y Superación de la Pobreza Urbana.

Finalmente, en 2003, Sedesol planteó una reorganización de sus programas que fue recogida por el PEF de ese año, consistente en la integración de sus acciones en sólo cinco programas y cinco programas con modalidades de operación,<sup>44</sup> además de las acciones y programas de sus entidades sectorizadas; lo que permite reducir la dispersión de esfuerzos y aumentar la concentración de los recursos para el logro de mayores impactos.

De esta forma, los programas que operan mediante modalidades son: Desarrollo Local (Microrregiones en 2002); Opciones Productivas; Desarrollo de los Pueblos y Comunidades Indígenas; Ahorro, Subsidio y Crédito para la Vivienda Progresiva Tu Casa; y Habitat.

Los que por su impacto social continúan funcionando por separado son: Incentivos Estatales, Empleo Temporal, Atención a Jornaleros Agrícolas, Jóvenes por México e Iniciativa Ciudadana 3×1.

También hubieron modificaciones en programas sectorizados, como en el caso de Indesol que subsumió todos sus programas en el de Coinversión Social.

## Comentarios finales

Este capítulo aportó información sobre los antecedentes y los principales planes y programas sociales vigentes en el sexenio, lo que servirá de marco para una mejor comprensión de sus evaluaciones. En la descripción realizada destaca una enorme dispersión de recursos en múltiples y pequeños programas, así como algunas duplicidades bastante evidentes.<sup>45</sup>

La estrategia Microrregiones cuenta con un diagnóstico basal que le permitirá en el futuro contar con indicadores cuantitativos de seguimiento a su gestión. Por el momento sólo se han desarrollado evaluaciones

<sup>44</sup>En 2002 fueron 13 los programas directamente a cargo de la Sedesol dentro del Ramo 20.

<sup>45</sup>Por ejemplo, en el caso de los programas Desarrollo de los Pueblos y Comunidades Indígenas de Ramo 20 y Desarrollo Económico Productivo de los Pueblos Indígenas del INI en 2002, lo que fue modificado posteriormente.



individuales de los programas revisados, mismas que se analizarán en los siguientes capítulos.

Su revisión deberá permitirnos, entre otras, aportar conclusiones sobre la viabilidad de medir sus contribuciones individuales a la disminución de la pobreza y la desigualdad o la conveniencia de evaluar su impacto a nivel de las grandes estrategias o políticas.

## Capítulo 8

# Primeros esfuerzos de evaluación de programas

### Introducción

PRESENTADAS LAS políticas y programas de desarrollo social aplicados en el país, corresponde exponer las particularidades que asume la evaluación aplicada a este tipo particular de la acción pública y revisar los primeros intentos por evaluar algunas de ellas.

Si bien se tiene conocimiento de algunos trabajos efectuados en el marco de los programas a cargo de Liconsa, Diconsa o Instituto Nacional de Solidaridad (Insol) en el periodo 1995-2000, no existen documentos públicos que permitan analizar esas experiencias. En cambio se cuenta con diversos trabajos de evaluación aplicados a Pronasol, y más aún a Progresá, que resultan vitales para entender el tipo de evaluación que se está realizando actualmente en Oportunidades, y en general, en los programas de desarrollo social.

### Particularidades de la evaluación de políticas sociales

A pesar de la importancia de los mecanismos de descentralización, externalización y focalización implementados en AL Cohen (2001) manifiesta que no existe claridad sobre los beneficios de su aplicación.<sup>46</sup> Su hipótesis es que esto se debe al traslado acrítico de estrategias de intervención que han sido exitosas en países europeos industrializados que intentan avanzar en formas de organización y gestión “posburocráticas”, cuando AL aún está permeada por todos los vicios “preburocráticos” y no existen condiciones que permitan recrear y regular los escenarios de mercado para el buen funcionamiento de la externalización. El Banco Mundial (2004a) también reconoce la falta de información sobre el impacto causado por los programas sociales en los países en desarrollo.

<sup>46</sup> Con base en un trabajo realizado en la Comisión Económica para América Latina (CEPAL), en 1998, plantea que los beneficios reales en términos de impacto y eficiencia derivados de la introducción de las dos primeras estrategias citadas en cinco programas sociales de la región latinoamericana, serían marginales o incluso nulos, cuando no se implementan correctamente.

Cohen (2001) resalta la importancia de proceder a su evaluación considerando:

1. la magnitud del gasto social;
2. el porcentaje del gasto social efectivamente redistributivo;
3. la eficiencia en la aplicación del gasto anterior; y
4. su impacto en la solución del problema.

Las razones de esta secuencia son obvias: no puede haber resultados relevantes en programas que carecen de recursos suficientes; no basta con gastar, pues muchos recursos pueden insumirse en los procesos sin llegar a la población objetivo;<sup>47</sup> y, suponiendo que lleguen, por lo general se desconoce su efectividad para resolver o aliviar el problema que ocasionó el programa, el grado en que justifican su costo y quiénes son sus verdaderos beneficiarios y perjudicados.

Considerando que los programas sociales se caracterizan por: requerir del mediano y el largo plazo para alcanzar sus objetivos redistributivos, por ser afectados por múltiples variables contextuales, una gran variedad de actores participantes con intereses en conflicto y una fuerte atribución de importancia a la participación social, a la transparencia y la responsabilidad de la gestión, al menos en el discurso (Cardozo, 2000b); su especificidad debe ser tenida en cuenta en su evaluación.

Nadie parece negar que la efectividad de un programa social depende también del grado en que éste recupera las necesidades sociales reales y responde a dichas circunstancias, lo que implica disponer de un diagnóstico que priorice los problemas a atender (Rossi y Freeman, 1993), que se apliquen mecanismos de control y evaluación desde el inicio, y que la información generada sea puesta a disposición del público para aumentar la transparencia y garantizar su participación (Franco, 2003). En el informe sobre la lucha contra la pobreza en México, el Banco Mundial (2004b) identifica como uno de los tres retos pendientes: fortalecer el monitoreo y la evaluación de impacto de los programas sociales.

<sup>47</sup> Se refiere un trabajo que demostraría en el caso chileno que sólo el 13 por ciento del gasto social total llega al 30 por ciento más pobre de la población; situación que se considera se repite en otros contextos nacionales de América Latina. Otro trabajo realizado en seis países latinoamericanos reporta que el 90 por ciento del gasto público social corresponde a gasto corriente.

En el caso del Progreso mexicano, que se revisará posteriormente, su evaluación calculó que los gastos administrativos o programáticos ascendían a 8.2 por ciento, los relacionados a la focalización del programa a 30 por ciento y los vinculados a su condicionalidad, 26 por ciento; los que fueron considerados porcentajes pequeños en comparación con programas mexicanos previos y otros dirigidos a grupos particulares en países extranjeros.

En el caso de la evaluación de la participación social, estrategia mayoritariamente considerada relevante en el marco de los programas sociales, Bichmann *et al.*, (1989) señalan que ésta puede ser estudiada tomando en cuenta diferentes áreas donde la misma puede incidir: orientación, liderazgo, organización, recursos y gestión. Cada uno de estos capítulos se califican y sus resultados se grafican con base en las respuestas obtenidas a un conjunto de preguntas aplicadas a los diferentes actores sociales.

En el presente trabajo se considera imprescindible la complementación de los agregados económicos planteados con indicadores cualitativos de pertinencia, coherencia, satisfacción, calidad, participación social, etcétera.

Además, es importante resaltar que la evaluación de políticas públicas es mucho más compleja que la de programas públicos, especialmente en el área social, ya que “sus metas son variadas, sus actores numerosos, los instrumentos diversificados, los resultados a menudo decepcionantes” (Kessler *et al.*, 1998, traducción libre). Según Perret (1993), algunos países se refieren a la evaluación en ambos campos (sin hacer diferencias entre ellos en Reino Unido y haciéndolas en EUA); otros sólo mencionan al campo de programas públicos (Canadá, Australia, Alemania); otros más, sólo a las políticas como conjunto de programas (Países Bajos), o bien, a la evaluación de políticas cuando en realidad se hace evaluación de programas (Francia). La situación mexicana se acerca bastante a esta última.

## Evaluaciones de Pronasol

Pronasol inició actividades en 1988, siendo el primer programa mexicano dirigido a combatir la pobreza en forma focalizada, privilegiando a las comunidades organizadas.

Entre los años 1990 y 1994, el programa fue objeto de múltiples intentos de evaluación parcial, ya sea a nivel estatal y/o enfocados a aspectos específicos del mismo. Pero, pocos fueron los casos en que se intentó su evaluación integral a nivel nacional. A continuación se exponen tres antecedentes relevantes en este último sentido:

a) Como ya se mencionó en el capítulo 2, Székely (1993) expresó la ausencia de elementos en la literatura que le permitieran responder a las preguntas formuladas por la OIT para proceder a su evaluación, salvo en lo relativo a los recursos ejercidos por el programa, su relación con el PIB y el gasto público. Las preguntas no respondidas se referían a: la identificación de efectos compensatorios y redistributivos, el nivel de focalización hacia los pobres, los efectos productivos en materia de creación de empleo y empresas,

en la construcción de infraestructura y dotación de servicios, y a la evaluación de la organización y operación institucionales. Concretamente, expresó que "...existen datos respecto a cantidades invertidas y/o volúmenes de obra, pero no evaluaciones propiamente dichas, en términos de costo beneficio, ni frente a metas numéricamente preestipuladas de reducción en los déficits que marcan los índices de pobreza..." (Székely, 1993: 32). Por esa razón, insistió en que su trabajo se basaba en su apreciación sensorial de la situación, enriquecida con el relato de algunas experiencias.

Székely procedió entonces a través del análisis y relación entre los tres componentes del programa: asistencial, participativo y competitivo.

En el primero resaltó el esfuerzo financiero del programa y lo comparó con los recursos que serían necesarios para superar las carencias de los sectores más pobres, tomando a la política de reactivación económica general como un dato.

En relación al componente participativo argumentó la fundamentación de la focalización hacia los pobres organizados, en lugar de hacerlo hacia los más pobres, con base en un criterio de tipo costo-beneficio. Además insistió en la ventaja de abandonar los canales burocráticos y corporativizados de circulación del gasto social.

En el tercero recurrió a la narración de experiencias de programas productivos en el estado de Oaxaca: Apoyo a Cafeticultores, Fondos Regionales de Solidaridad para el Desarrollo de los Pueblos Indígenas, Fondo para el Desarrollo Pesquero del Litoral Oaxaqueño y Fondo Nacional de Empresas de Solidaridad.

En todos los casos retomó los datos numéricos disponibles en tanto aproximaciones para apoyar sus reflexiones cualitativas. También puso especial énfasis en analizar las críticas que se le hacían al programa.

Concluyó en la urgencia de ampliar el programa mediante un esfuerzo económico y político sin precedentes, acompañado de una modificación de las funciones ejercidas por las instituciones y las bases sociales organizadas.

El intento de evaluar el Pronasol en su conjunto dejó clara la insuficiente información disponible para la realización de un trabajo riguroso que pretendiera basarse en el seguimiento de metas y medición de impacto.

*b)* Un segundo intento fue realizado por Jaso y Merino (1994), quienes procedieron a realizar una síntesis evaluativa y crítica de los trabajos parciales realizados, cuyas conclusiones resultaban a menudo complementarias y en otros casos definitivamente contradictorias.

Consideraron al programa como una política pública con propósitos definidos, que ameritaba una evaluación lo más objetiva posible, al mar-

gen de los intereses políticos del gobierno y los partidos. Esto implicaba dificultades importantes debido a: el limitado avance de la evaluación de políticas públicas, la complejidad y novedad de la organización del programa, la importancia atribuida a la organización social, la diversidad de métodos empleados en las evaluaciones parciales analizadas, las diferentes fechas en que fueron realizadas, etcétera. Resolvieron entonces ofrecer una visión, de énfasis cualitativo, que recogiera todas las opiniones con su respectivos fundamentos y evidencias, tratando de rescatar sus ideas centrales.

Dichas ideas indujeron la estructura final de su trabajo: dos grandes capítulos sobre la operación, el programa y la organización social y otros dos complementarios sobre su impacto económico y político.

El primero revisa los objetivos del programa, su organización, mecanismos de implementación y resultados verificables, señalando aciertos y deficiencias. Concluye que las evaluaciones coinciden en destacar la complejidad ejecutiva del Pronasol y su operación centralista. Agregan que los recursos no siempre fueron destinados con precisión a los grupos más pobres, ni a los proyectos que podrían ayudarlos más.

El capítulo sobre organización social, considerada la diferencia específica de Pronasol, analiza la integración y funcionamiento de los comités de solidaridad que aportaron dinero y mano de obra al programa, el perfil socioeconómico de los participantes, su nivel de satisfacción y expectativas de continuidad. Se concluye que la misma impulsó una relación diferente entre el gobierno y la sociedad, potenciando formas tradicionales de cooperación comunitaria, entre las que sobresalen los resultados de participación obtenidos por los comités de Escuela Digna.

En materia de impactos económicos el trabajo discute la confiabilidad de los datos INEGI-ONU-CEPAL que referían una disminución de la pobreza externa de 14.9 a 13.6 millones de personas entre 1989 y 1992. Concluye que el programa fue eficaz en el mejoramiento de la calidad de vida de los grupos más pobres cuando ésta se mide en términos de acceso a satisfactores colectivos, pero ha sido insuficiente para garantizar los ingresos familiares necesarios para la satisfacción de sus necesidades esenciales.

La crítica más frecuente al programa es que habría sido utilizado para favorecer electoralmente al PRI, pero los aspectos políticos del mismo fueron los más debatidos y con menor información empírica disponible. Sin embargo, el trabajo afirma que los datos existentes sobre distribución del gasto no muestran evidencias de que Solidaridad haya reorientado la misma regionalmente en función de las elecciones locales.

Se concluye resaltando los aciertos del programa pero también su insuficiencia para revertir las inercias de funcionamiento y las causas estructurales de la pobreza.

c) En tercer lugar, con la intención de contribuir al desarrollo de una metodología específica para la evaluación del impacto socioeconómico del programa, el Colegio de México realizó un trabajo coordinado por Reyna y Rocha (s/f), en el que también participó Székely como uno de los investigadores del equipo. El mismo constituye un antecedente obligado de lo que serían las posteriores evaluaciones de programas sociales.

Dicho trabajo define ante todo un marco de referencia para la evaluación, comenzando por identificar las hipótesis subyacentes en el programa: la población objetivo llegará a alcanzar mayores ingresos al mejorar su estado de salud, adquirir mayores conocimientos, contar con mayor infraestructura de servicios y aumentar o mejorar su capacidad de participación en la toma de decisiones. Sin embargo, esta formulación implica aún otra hipótesis que el documento no explicita: que la obtención de mayores ingresos redundará en mejores condiciones de vida para la población (objetivo final, señalado en primer lugar, por el programa).

El documento sugiere tres aspectos interdependientes a ser considerados en la evaluación: la adecuación del programa a las necesidades y demandas de la población, la organización interna del programa y la relación con su población objetivo, y el impacto en la generación de un ingreso adicional para la última.

Estos aspectos y la congruencia entre el objetivo final y las estrategias del programa, indicada en el punto inicial, forman parte de los temas de diseño, funcionamiento y medición de impactos que esta tesis considera imprescindibles, pero poco considerados en los términos de referencia e informes de evaluación de 2002, que resaltan más los recursos aplicados, las metas alcanzadas, la cobertura, los resultados logrados, y el apego de la operación a la normatividad.

Reyna y Rocha reconocen la heterogeneidad de estrategias del programa, los distintos tiempos de avance de los proyectos y las dificultades para identificar tanto los impactos esperados como los imprevistos, los inmediatos como los mediatos, por lo que sólo proponen una metodología de “aproximación” a los impactos del Pronasol por estados, basada en dos dimensiones de análisis: los objetivos propios de cada estrategia y las características organizativas y de participación de la población.

En los programas sociales subsiguientes se abandonará este énfasis de Pronasol en la organización de la población; sin embargo, el discurso de las

políticas seguirá insistiendo en la participación de la misma, aunque descuidando los canales para su implementación y evaluación.

La propuesta de metodología de evaluación analizada sugiere la consideración de diversos conjuntos de indicadores: de medición operativa de los niveles de pobreza extrema, comunes a todos los componentes, y específicos para cada uno según sus campos de acción (salud, educación, infraestructura, etcétera).

Actualmente, casi todos los programas se limitan a la medición específica de sus objetivos y metas inmediatos. La medición de la evolución de la pobreza se hace por otros medios (encuestas del INEGI y estudios que utilizan su información) y resulta prácticamente imposible saber si todos los programas aplicados han contribuido a la misma y en qué proporción, por lo que suele correlacionarse sólo con el mayor de los programas (Oportunidades).<sup>48</sup>

El documento del Colmex plantea el conjunto de preguntas a responder, ofrece una propuesta de indicadores y señala las relaciones entre los mismos, sugiere la forma de evaluar la participación de cada actor relevante del programa, la manera de determinar la muestra estadística para la aplicación de encuestas, y la forma de presentar los resultados.

El capítulo relativo a la medición del cambio en los niveles de pobreza es prácticamente indiscutible desde el punto de vista teórico: se trata de calcular un índice antes y después de las acciones de Progresá y medir la relación costo-beneficio como el cociente entre el costo total de las acciones y el cambio logrado en el nivel de bienestar. Sin embargo, las seis páginas de extensión del capítulo no abordan la complejidad que implica la identificación y el cálculo del índice de pobreza y tampoco plantean el riesgo de adjudicar al programa efectos provocados por otras variables y los métodos disponibles para su reducción (experimentación con grupo de control).

Otro breve capítulo de gran interés lo constituye el estudio de la relación entre los indicadores de impacto y la participación. Actualmente se cuenta con pocos casos de participación significativa de los beneficiarios pero, para no promover o rechazar la misma sin evidencia empírica, debería diseñarse un experimento controlado para determinar si la misma tiene o no efectos positivos y en qué condiciones.

La presentación de resultados se limita, en dos cuartillas, a indicar que se dará respuesta a las preguntas planteadas.

<sup>48</sup>El documento del BM (2004b), en cambio, plantea que no se sabe en qué porcentaje la pobreza ha disminuido por los programas gubernamentales aplicados y por las remesas recibidas por los mexicanos desde el exterior.



Como puede verse, el documento no profundiza en las grandes dificultades que en la operativización de la metodología de evaluación se plantean, pero constituye un primer esbozo de cómo intentarla que se acompaña de valiosos anexos (indicadores, cuestionarios, ejemplos de medición de pobreza y de cambios en el nivel de bienestar, etcétera) y señala aspectos problemáticos que aún hoy no han sido plenamente resueltos.

### Evaluaciones de Progresá 1998-2000

El Programa de Educación, Salud y Alimentación (Progresá) constituye el antecedente inmediato del actual Programa de Desarrollo Humano Oportunidades. Estuvo bajo la responsabilidad de la Sedesol, y contó desde 1997 con una Coordinación Nacional (ConProgresá) integrada por los titulares de la secretaría anterior, quien lo presidía, y de la SSA, SEP y SHCP.

Dicho programa incluyó tanto transferencias de efectivo a las madres de familia para mejorar los niveles de consumo de los pobres en el corto plazo, como inversiones en capital humano para lograr su desarrollo social, mismas que dieron inicio en mayo de 1998. Mientras Pronasol, su antecesor, había puesto el énfasis en la organización de las comunidades, Progresá se diseñó para dirigirse a individuos y, después de su discusión, se resolvió reorientarlo hacia las familias. De esta forma el programa procedió con una doble focalización: a las localidades más marginadas y, dentro de ellas, a los hogares más pobres. Este diseño implicó que los hogares muy pobres ubicados en localidades no marginadas quedaran fuera de las metas del programa.

Fue el primer programa mexicano (y tal vez el único) cuya evaluación de impacto fue prevista desde su diseño,<sup>49</sup> partiendo de un diagnóstico de la situación antes de su inicio, medida a través de la aplicación de encuestas a hogares en octubre-noviembre de 1997 y marzo de 1998; y aplicando tres evaluaciones posteriores en octubre-noviembre de 1998, junio y noviembre-diciembre de 1999. La misma revisó los impactos en materia de educación, nutrición, salud y pobreza rural, así como la operación general del programa. Posteriormente se le hicieron ajustes y se continuó con sus evaluaciones, las que en 2003 habían generado 11 volúmenes de documentos.

Si bien en la evaluación inicial participaron organizaciones mexicanas como la UIA, el INSP y la Escuela Nacional de Antropología e Historia (ENAH), la responsabilidad del proceso estuvo a cargo del International Food Policy Research Institute (IFPRI) de Washington, D.C., mediante contratación

<sup>49</sup> En esto habría influido en parte la presión del Banco Mundial.

directa y bajo la dirección del doctor Emmanuel Skoufias, quien actualmente forma parte del BID.<sup>50</sup> O sea que el gobierno no le confió la responsabilidad de esta primera evaluación a los propios connacionales, sino que recurrió al empleo de expertos de un instituto internacional de investigación.

El gobierno del licenciado Ernesto Zedillo pagó por esta evaluación 4.35 millones de pesos en 1998, 11.31 en 1999 y 12.30 en 2000. Este monto no cubría el trabajo de campo, del que se encargaba el mismo programa.

La evaluación del Progreso utilizó un diseño experimental con la finalidad de garantizar un alto rigor en la identificación de relaciones causales entre el programa y los impactos identificados, con mediciones “antes-después”, talleres en los siete estados piloto y discusión en grupos focalizados (médicos, enfermeras, promotoras, profesores, etcétera), además del empleo de técnicas tradicionales como encuestas y entrevistas.

La evaluación realizada destaca, a nivel del diseño del programa, una focalización más clara de lo que sucedía habitualmente en programas similares de América Latina, una importante sinergia debida a la consideración de tres sectores-clave (salud, educación y alimentación), la utilización de transferencias condicionadas a la participación de las familias beneficiadas en escuelas y centros de salud, y el apoyo complementario a los últimos.

Para Skoufias (2003), los elementos-clave que garantizan una evaluación exitosa y rigurosa son: que la evaluación sea un componente programado desde el diseño del programa, que tenga objetivos claros, que se cuente con el apoyo del ámbito institucional responsable del mismo, que el diseño de la evaluación ofrezca mediciones creíbles, que cuente con un adecuado marco lógico, presupuesto suficiente y, de preferencia, un diagnóstico previo a la implantación del programa.

El diseño de la evaluación de Progreso hubiera requerido idealmente una selección aleatoria de beneficiarios y no beneficiarios (grupo de control), a nivel individual, pero debido a que la rigurosidad es costosa, se optó por una selección aleatoria a nivel de localidades, que no garantiza de igual manera la ausencia de sesgos en la medición de impactos.<sup>51</sup> Las localidades seleccionadas fueron 506, compuestas por 302 de un total de 4,546 en que el programa fue aplicado en su fase II y 186 de 1,850 en que esto no se hizo (grupo de

<sup>50</sup> Consultar al respecto Skoufias (2001 y 2003), en [www.ifpri.org](http://www.ifpri.org). También los ocho tomos publicados en el año 2000 con las evaluaciones realizadas en 1998-2000.

<sup>51</sup> Consultar en [www.ifpri.org](http://www.ifpri.org). Behrman y Todd (febrero de 1999), en su reporte a Progreso “Randomness in the experimental samples of Progreso”, en el que manifiestan que no hay diferencias significativas en las características generales de edad, educación, ingreso o acceso a la salud entre el grupo de tratamiento y el de control; sin embargo, a nivel individual y familiar, encuentran diferencias en algunas variables que, aunque en promedio son muy pequeñas, el gran tamaño de la muestra, hace que resulten significativas a niveles convencionales (10 por ciento en su análisis).

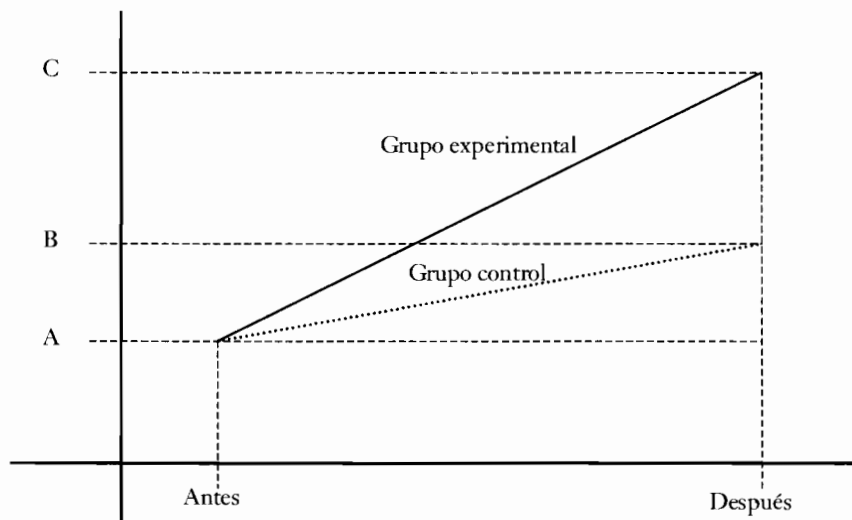
control), distribuidas en los siete estados piloto en que se inició la aplicación del programa: Guerrero, Hidalgo, Michoacán, Puebla, Querétaro, San Luis Potosí y Veracruz. Incluyó a 24,077 hogares y más de 100,000 personas, o sea, aproximadamente el uno por ciento de los hogares apoyados.

Skoufias señala que este importante tamaño muestral, podría ser incluso insuficiente para medir el impacto en algunas áreas; lo que de inmediato refuerza las dudas que dejan las evaluaciones de otros programas sociales en que la muestra dista mucho de alcanzar este tamaño, independientemente de los problemas enfrentados para garantizar una selección aleatoria.

El diseño experimental tiene la ventaja de permitir observar la diferencia entre las mediciones “antes-después” del programa (C-A en la gráfica 1), comparada con la evolución sufrida por grupos de similares características iniciales (punto de partida aproximadamente coincidente en la gráfica) a los que el programa no se aplicó pero que, por otro tipo de razones ajenas al programa mismo, podrían también haber evolucionado, favorable o desfavorablemente (B-A), lo que a su vez permite identificar el verdadero impacto provocado por las acciones evaluadas (C-B en la gráfica).

Entre los resultados globales obtenidos de esta forma para Progresá, resaltan:

GRÁFICA 1  
MEDICIÓN DE IMPACTO DEL PROGRAMA



1. Muy buena focalización geográfica en zonas rurales.
2. Bastante precisión en el método de selección de hogares.
3. Fuerte impacto en el ingreso real de las familias.
4. Ausencia de impactos negativos como disminución del tiempo dedicado por los padres de familia al trabajo.
5. Errores administrativos menores como demora en la distribución de las transferencias, errores en listas de beneficiarios, etcétera.
6. Efectos indirectos positivos en las localidades con familias apoyadas debido a la derrama promedio de 30,000 pesos mensuales adicionales.

A pesar de la rigurosidad del diseño evaluativo, el lenguaje empleado en la presentación de resultados transmite la conciencia de los evaluadores sobre el grado de riesgo implícito en algunos de ellos: "...los resultados de la investigación sugieren...", "Los datos sugieren que Progresa..." o "Parece que Progresa..." (IFPRI, 2002: 3, 5 y 7, respectivamente). Skoufias insiste en su insuficiencia para determinar algunos impactos que serían deseables, como:

1. ¿Cuál es el monto mínimo a transferir como beca para lograr un impacto real en la permanencia de los niños en la escuela y su acceso regular a los centros de salud?
2. Además de la permanencia en la escuela, ¿realmente se incrementó la calidad de la educación y el aprendizaje logrado por los estudiantes? La evaluación no encontró indicios de esto último, que constituiría la misión del programa educativo.
3. ¿El programa puede constituirse en una red de seguridad social en momentos de crisis?
4. ¿Cuánto tiempo puede continuar un programa como éste, que requiere de un monto importante de recursos económicos, si no demuestra claramente su eficacia?

Estas fueron algunas de las dudas que quedaron planteadas en las evaluaciones efectuadas por el IFPRI.

Durante el año 2000 Progresa atendió a 2.478 millones de hogares. Se continuó con la aplicación de las encuestas semestrales a hogares, en dos nuevas rondas (en mayo y noviembre) para completar los tres años de seguimiento, mismo que había dado inicio en marzo de 1998. La novedad fue que, en estas fechas, el grupo de control había sido incorporado al programa, con lo que el diseño experimental original sufrió un cambio: ya no había la contundente posibilidad de comparar resultados entre beneficiarios y no benefi-

ciarios, sino sólo entre dos grupos de beneficiarios, incorporados al programa con menos de dos años de diferencia (mayo de 1998 y enero de 2000). Las razones del cambio, que metodológicamente dificultaron su evaluación, fueron de carácter ético: imposibilidad de continuar negando al grupo de control su participación en los beneficios generados por el programa, siendo que los recursos eran suficientes para su inclusión.

### Auditorías de desempeño de Progresá 1999-2000

En forma paralela al proceso descrito, a partir de 1999, la ASF comenzó a realizar las denominadas “auditorías de desempeño” a los programas federales. Entre los programas evaluados incluyó al Progresá: su componente educativo en 1999<sup>52</sup> y su componente de salud en 2000. Los mismos fueron analizados como parte de la responsabilidad que corresponde a la SEP y SSA, no a la Sedesol ([www.asf.gob.mx](http://www.asf.gob.mx)). Dentro de esta última los programas seleccionados fueron Jornaleros Agrícolas y Apoyo a las Empresas en Solidaridad (Fonaes).

Los informes de dichas auditorías comienzan por fundamentar las razones que llevaron a la selección del programa, se apoyan en información documental del mismo y realizan un trabajo de campo específico para el cual solicitan el apoyo de otras instituciones (UNAM). También retoman los informes de los estudios realizados por el IFPRI, el BM y la CEPAL.

Su trabajo utiliza un conjunto de indicadores para el análisis de cada programa, en el que figura:

- congruencia del objeto social del programa con las prioridades nacionales y sectoriales de mediano plazo (el sexenio en curso);
- eficacia en el cumplimiento de las metas programadas;
- cobertura respecto del universo objeto del programa;
- eficiencia en la operación y aplicación del presupuesto; e
- impacto del programa.

La auditoría de desempeño aplicada a Progresá<sup>53</sup> en el año 2000 realiza observaciones y promueve acciones específicas, entre las que resalta la insistencia en que el programa no había establecido indicadores de impacto

<sup>52</sup>No fue posible acceder a dicho documento debido a que sólo están disponibles en Internet los informes de la Cuenta Pública del año 2000 en adelante.

<sup>53</sup>Sorprende que en diversas partes del documento se hable de Oportunidades (aún no creado en el año 2000) en lugar de referirse a Progresá. Consultar, por ejemplo, las acciones promovidas 00-20G00-7-373-07-001, 00-20G00-7-373-07-002 y 00-12100-7-373-07-008.

que permitieran evaluar el cumplimiento de su objetivo (mejorar el nivel de bienestar y calidad de vida de la población beneficiaria) en su programación anual y de mediano plazo. Entre las recomendaciones de menor magnitud se encuentran: implementar un sistema de quejas y denuncias, realizar campañas de difusión del componente de salud, controlar que el personal médico y paramédico no exija ninguna aportación por la entrega del suplemento alimenticio, etcétera.

### Evaluación y auditoría de desempeño de Progres 2001

En 2001 se pusieron en marcha dos nuevas acciones: la ampliación de la cobertura del programa hacia las zonas urbanas, y la entrega de becas educativas a estudiantes de nivel medio superior.

De esta forma se incorporaron al programa, originalmente diseñado sólo para el medio rural, 600 localidades urbanas y 1.76 millones de nuevas familias, o sea que el número de familias atendidas aumentó en un 72 por ciento. Para el ciclo escolar 2001-2002, según datos de Sedesol, se logró la inscripción de 3.3 millones de becarios en los niveles de primaria, secundaria y educación media superior, cifra superior en 34 por ciento al número de niños y jóvenes atendidos en el ciclo escolar 2000-2001. En particular se entregaron becas educativas para continuar estudios de nivel medio superior a 267,733 jóvenes que vivían en condiciones de pobreza extrema.

Por otra parte, para corregir errores de exclusión en el proceso de selección de familias se estableció un mecanismo de densificación de la cobertura en las localidades rurales ya atendidas. Con el fin de atender a los beneficiarios del programa en las ciudades se instalaron 196 módulos que operan bajo un esquema automatizado de atención, basado en el uso de tecnologías informáticas y de transmisión electrónica de datos. Cada módulo fue habilitado con equipo de cómputo y una plataforma de cartografía digitalizada, lo que permitió reducir el tiempo desde el levantamiento de la información socioeconómica de las familias hasta su incorporación.

El programa le solicitó a una institución nacional que propusiera los instrumentos de línea basal urbana en pequeñas localidades. Dicha línea fue diseñada entre mayo y septiembre con grupo de tratamiento y de control y se inició el trabajo de campo. A finales de 2001, al realizarse las modificaciones ya mencionadas al programa, se abandonó el seguimiento de los paneles y se decidió la ampliación del programa a ciudades de mayor tamaño, sin respetar el grupo de control diseñado.

Estrictamente, la evaluación de la gestión fiscal 2001, a diferencia de las anteriores que duraron dos años y medio, fue realizada en un mes. La misma incluyó: una evaluación del funcionamiento del programa a nivel estatal, basada en el análisis de los indicadores de seguimiento, gestión y evaluación de 1998 al 2001; y, un análisis de los impactos del programa basados en las rondas recientes de la Encuesta de Evaluación en Localidades (Encel), no utilizadas en las evaluaciones previas. O sea que no incluyó levantamiento específico de información en campo. Su costo fue de 16.13 millones de pesos.

La evaluación identifica problemas relevantes en el mencionado sistema de indicadores (respecto de la continuidad de los incluidos, de su desglose, y de su formato de presentación, algunas redundancias y una clasificación no es siempre clara, algunos indicadores cuya relevancia para la evaluación del programa es limitada, y otros más relevantes que están ausentes) y dificultades para hacer una evaluación a nivel estatal con datos agregados, por lo que señala que sus resultados son indicativos y ameritan estudios en mayor profundidad. Entre ellos resaltan:

- Una distribución de las transferencias monetarias por estado congruente con sus niveles de pobreza monetaria y las necesidades relevantes de salud, lo cual no resulta igualmente claro en materia educativa.
- Avances notables en el cumplimiento de las condiciones de corresponsabilidad que impone el programa a sus beneficiarios en materia de educación y salud.
- Avances claros en relación con las tasas de deserción escolar y de desnutrición infantil.
- Persistencia de altas tasas de repetición de los beneficiarios y de una brecha importante entre éstos y los promedios estatales en materia educativa ([www.progresagob.mx](http://www.progresagob.mx)).

Por su parte, las nuevas rondas de Encel dieron lugar a dos estudios de medición de impacto: un análisis del comportamiento del grupo de control, al ser incorporado al programa; y una evaluación de algunos resultados de mediano plazo (sólo tres años) en los jóvenes beneficiarios del mismo. Los últimos incluyeron el cálculo de la probabilidad de que los estudiantes becados por Progresá en el nivel básico continuaran con su educación media superior (aún no se otorgaban becas para ese nivel), o, si se incorporaban al mercado de trabajo, obtuvieran mejores empleos y remuneraciones, así como su relación con la migración, y la propensión de los niños a enfermar.

De los resultados obtenidos en los anteriores, se destaca:

- Progresá benefició al 40 por ciento de las familias en las zonas rurales.
- El grupo de control se comportó en forma muy similar al de tratamiento, al momento de su inclusión al Progresá, en aquellas variables en que se puede identificar un efecto inmediato, como inscripción escolar, visitas a centros de salud y gasto de los hogares.
- Se confirmaron también efectos en los años de escolaridad de los niños y niñas apoyados por el programa: 0.15 y 0.38 años más, respectivamente, que los niños que no recibieron becas.
- Los datos "sugieren" un pequeño efecto en la inscripción de los ex becarios a la educación media superior, que se esperaba se incrementara con el tiempo, lo que se reforzaría por un nuevo elemento: el otorgamiento de becas de Progresá a este nivel a partir de 2001.
- La información disponible "no sugiere" que los jóvenes de las localidades donde opera Progresá tengan mayor probabilidad de migrar fuera de sus comunidades, incluso cuando han dejado de ser elegibles para la obtención de becas de estudio.
- Resulta muy prematuro evaluar los efectos del programa en las oportunidades de empleo e ingresos de los ex becarios, ya que la mayoría aún continúa en la escuela, siendo los que la abandonaron quienes no lograron mejorar su desempeño a pesar de la beca ([www.progresá.gob.mx](http://www.progresá.gob.mx)).

Como ya se comentó, metodológicamente hablando, el diseño evaluativo se debilitó con la inclusión del grupo de control al programa, por lo que se recomendó la utilización de una estrategia no experimental de estimación<sup>54</sup> para la medición de impactos del programa en el largo plazo.

Ese mismo año se presentaron resultados de una evaluación cualitativa realizada en zonas semiurbanas. Con estos dos trabajos, se cubrió la evaluación anual de 2001, si bien los resultados se basaban en información recogida mediante las Encel de mayo y noviembre del 2000.<sup>55</sup>

Al igual que el año anterior, la ASF llevó a cabo la auditoría de desempeño de programas a cargo de la Sedesol (Ahorro y Subsidios Para la Vivienda,

<sup>54</sup>Sugiere la estrategia de *matching*, emparejamiento o apareamiento, a revisar en J.J., Heckman, R. Lalonde y J. Smith (1999), "The economics and econometrics of the active labor market program", en *Handbook of labor economics*, vol. 3A, ed. O. The Netherlands, Nort Holland, Ashenfelter y D. Card Amsterdam, .

<sup>55</sup>En realidad las evaluaciones de Progresá-Oportunidades nunca corresponden estrictamente al año calendario, como sucede con los otros programas. Incluso la evaluación 1998-2000, utilizó las bases de datos disponibles hasta finales de 1999.



Crédito a la Palabra, Coinversión Social y Empleo Temporal) e incluyó nuevamente al Progresá en su componente alimenticio (dentro del sector salud).

En esta ocasi3n su informe reconoce que el programa cuenta con los indicadores y evaluaciones necesarias para medir el cumplimiento de sus objetivos y metas, así como sus procesos operativos. Recomienda dar cabal cumplimiento a las normas que rigen la rendici3n de cuentas, reforzar mecanismos de control y supervisi3n en las unidades m3dicas relativas al cumplimiento de las acciones que deben realizar las familias beneficiarias como parte de su corresponsabilidad, etc3tera.

Cabe destacar tambi3n que Progresá fue uno de los pocos programas sociales que fue evaluado por ONG. En 1999, y como parte del seguimiento a los acuerdos de Beijing, la campaa “El Banco Mundial en la Mira de las Mujeres” analiz3 un conjunto de programas de AL que, en el caso de M3xico, incluy3 al Programa de Ampliaci3n de Cobertura en Salud (PAC) y al Progresá (Frade, 2000). La evaluaci3n relativa a este 3ltimo se diseaó con base en una muestra de 309 beneficiarias y 123 no beneficiarias, todas pertenecientes a organizaciones sociales y distribuidas en los estados de Chiapas, Campeche, Guanajuato, Jalisco, Oaxaca, Puebla, Yucatán y Quintana Roo, a las que se aplic3 una encuesta durante el segundo semestre de 1999. Adem3s se realizaron 83 entrevistas a maestros, promotoras del programa, personal de centros de salud, funcionarios federales, estatales y municipales, así como organizaciones sociales.

Entre sus resultados reconoce la asignaci3n de recursos gubernamentales, los costos de operaci3n bajos del programa (aunque basándose en el trabajo gratuito de las mujeres) y el diseaó de un componente de g3nero. Sin embargo, su juicio es muy negativo en cuanto a la ausencia de participaci3n corresponsable de las beneficiarias en la toma de decisiones, seguimiento y evaluaci3n del programa, desigualdad y falta de criterios para la asignaci3n de apoyos, exclusi3n de las zonas rurales de mayor pobreza por falta de infraestructura educativa y de salud, cobro de cuotas importantes por servicios supuestamente gratuitos, no consideraci3n del nivel de desempleo en el pa3s y de las escasas oportunidades existentes en el medio rural, trato indigno de las beneficiarias como indigentes, generaci3n de divisiones entre la poblaci3n, constituci3n de un mecanismo de control social, falta de adecuada conceptualizaci3n de la perspectiva de g3nero, ignorancia de problemas relevantes en ese campo, etc3tera. Resalta la forma en que se refieren al programa algunas entrevistadas del estado de Puebla, “dinero del diablo”, porque “...los recursos no les alcanzan cuando llegan, «no rinden», pero, adem3s, no se los ganaron” (Dora Ramos y compaaeras, en Frade, 2000:

86). Como consecuencia de los múltiples problemas identificados, el informe concluye que Progresá no permite superar la pobreza ni mejora la posición de las mujeres.

Finalmente, y con una opinión totalmente contrapuesta a la previa, el propio BM (2004a) considera que México ha sido pionero por la aplicación de evaluaciones de impacto de Progresá, las que permitieron al gobierno mexicano contar con evidencias de su éxito y fomentaron la creación de programas similares en otros países de AL (Honduras y Nicaragua con diseños experimentales y Colombia y Brasil con cuasiexperimentos)

### Comentarios finales

Los documentos revisados permiten aquilatar la insuficiente información y disponibilidad de conocimientos especializados de evaluación de programas sociales a inicios de los noventa y el importante avance logrado hasta el año 2000, gracias a las evaluaciones realizadas a Progresá por el IFPRI y las auditorías efectuadas por la ASF.

Con base en estos antecedentes, la Cámara de Diputados estableció un parteaguas en la situación al decidir la ampliación del proceso de evaluación externa a la gestión de 2001 de todos los programas que reciben recursos fiscales y se hayan sujetos a reglas de operación. Los capítulos restantes analizan los procesos seguidos y los resultados alcanzados, comenzando en el siguiente por la descripción del marco jurídico aplicable a la realización de evaluaciones.

## Capítulo 9

# Marco jurídico de la evaluación de programas

### Introducción

PARA FACILITAR la comprensión de las evaluaciones que serán analizadas en los siguientes capítulos, en el presente se revisa la normatividad que rige la organización y planeación formales de la actual evaluación de programas sociales, las instancias participantes y la descripción selectiva de sus funciones vinculadas a evaluación.

Específicamente, se describe el marco jurídico vigente para las evaluaciones externas aplicadas a partir del año 2001.

### Control y evaluación al interior de la administración pública federal

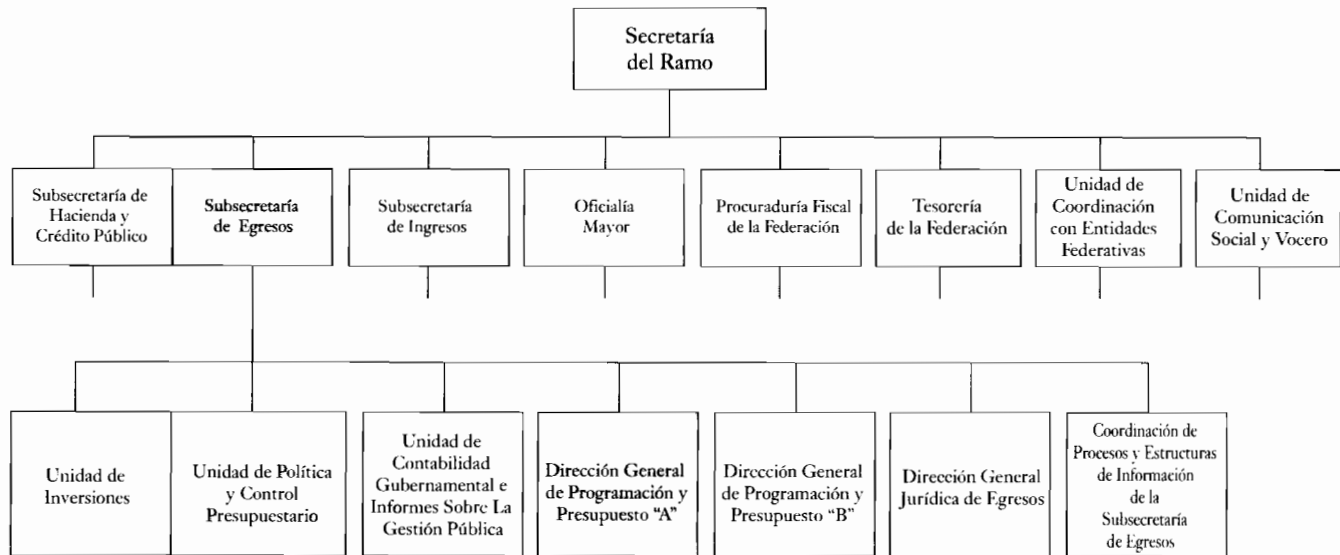
La Ley Orgánica de la Administración Pública Federal establece las competencias de las diversas secretarías de Estado vinculadas a la evaluación, sus presupuestos y la utilización de sus resultados (véase anexo 2).

De acuerdo con su artículo 31, corresponde a la Secretaría de Hacienda y Crédito Público (SHCP), preparar el proyecto de Presupuesto de Egresos de la Federación (PEF), a través de la Subsecretaría de Egresos, que incluye, en la Dirección General de Programación y Presupuesto “B” (véase organigrama 1), el de las instancias operadoras de programas sociales, dentro del cual se encuentra, a su vez, el destinado a la realización de sus evaluaciones externas. Al mismo tiempo supervisa la aplicación del gasto público federal en la Unidad de Política y Control Presupuestario.

Al ser entrevistada, su directora general adjunta de Desarrollo Social, Economía y Trabajo manifestó que ella revisa personalmente los informes de las evaluaciones externas, no teniendo otra ingerencia en el proceso. Agregó que desde el año 2000, la SHCP determina el techo financiero de cada secretaría, Sedesol incluida, se lo comunica y las secretarías lo distribuyen con total autonomía. Luego sólo les corresponde realizar el control presupuestal.

ORGANIGRAMA I

SECRETARÍA DE HACIENDA Y CRÉDITO PÚBLICO  
(Parcial)



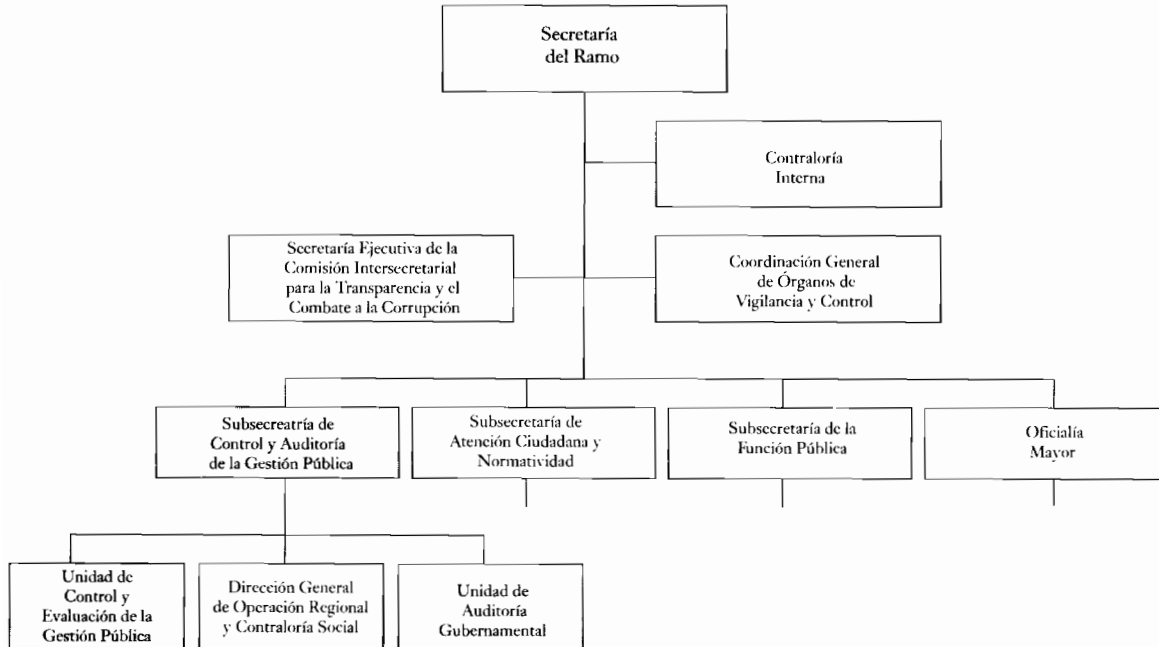
En su artículo 37, la citada ley establece que a la Secretaría de la Función Pública (SFP) le corresponde la vigilancia y el control del cumplimiento de normas y compromisos administrativos, así como la posible realización de auditorías y evaluaciones a las dependencias y entidades de la administración pública federal, para garantizar la calidad de los programas y propiciar la transparencia y la rendición de cuentas a la ciudadanía, lo que ejecuta a través de la Subsecretaría de Control y Auditoría de la Gestión Pública y las unidades de control y evaluación de la gestión pública y de auditoría Gubernamental (véase organigrama 2).

En materia de evaluación le corresponde también vigilar el cumplimiento del Acuerdo firmado por la SHCP y la entonces Secodam (en adelante, el Acuerdo), publicado en el *Diario Oficial de la Federación* (30 de abril de 2002, segunda sección: 4-7), con la intención de que las evaluaciones sean de calidad para mejorar el desempeño de los programas federales, se generen insumos adicionales a las dependencias para mejorar el diseño y resultados de sus programas, se aporten insumos sobre las deficiencias y bondades de las evaluaciones externas a los coordinadores de sector, se ofrezcan elementos para la preparación del proyecto de presupuesto al representante del Ejecutivo, se brinde apoyo técnico a la Cámara de Diputados sobre el avance y valoración de los programas federales, y que se pueda posicionar a la Dirección General de Operación Regional y Contraloría Social de la SFP como un ente capaz de emitir un dictamen confiable sobre la evaluación externa de los programas federales y que emita recomendaciones de política pública en función del análisis de sus resultados (Alamilla, s/f).

El director de Desarrollo Organizacional y Profesionalización, al ser entrevistado, coincidió en que sus principales funciones al respecto son verificar el cumplimiento de los requisitos del acuerdo y la confiabilidad de sus resultados, y agregó una tercera: concentrar los principales resultados y conclusiones de los informes que tengan sustento e investigar puntos recurrentes que puedan agregar valor al trabajo de SFP. Informó también que la Unidad de Control y Evaluación de la Gestión Pública realiza el análisis de los indicadores de los Programas, mientras la Dirección General de Operación Regional y Contraloría Social los analiza con la intención de incrementar la eficiencia del proceso de evaluación externa a los programas.

El actual Sistema Nacional de Control y Evaluación de la Gestión Pública a cargo de la SFP tiene como objetivos promover y coadyuvar al uso eficaz y eficiente de los recursos públicos, el desarrollo administrativo, la calidad de los servicios públicos, y el combate a la impunidad y la corrupción.

ORGANIGRAMA 2  
SECRETARÍA DE FUNCIÓN PÚBLICA  
(Parcial)



El sistema está constituido por el conjunto de instrumentos jurídicos, instancias, esquemas y mecanismos para el control y la evaluación de la gestión pública y la colaboración en materia de desarrollo administrativo. En él participan el gobierno federal, los gobiernos de los estados y del Distrito Federal.

En el marco del anterior funciona la Comisión Permanente de Contralores Estados-Federación, que facilita la vinculación entre los ámbitos señalados con objeto de intercambiar experiencias y realizar acciones en mutuo apoyo.

Méndez (2000) propone que se dote de mayor fuerza, capacidad y autonomía al sistema de seguimiento, control y evaluación, dotando a la SFP de un estatuto autónomo, similar al del Banco de México.

En dos numerales de su artículo 32, la Ley Orgánica de la Administración Pública Federal determina las competencias de Secretaría de Desarrollo Social (Sedesol) en materia de evaluación:

- I- Formular, conducir y evaluar la política general de desarrollo social para el combate efectivo a la pobreza; en particular, la de asentamientos humanos, desarrollo urbano y vivienda; y
- V- evaluar la aplicación de las transferencias de fondos en favor de estados y municipios, y de los sectores social y privado que se deriven de las acciones e inversiones convenidas, en los términos de las fracciones anteriores.

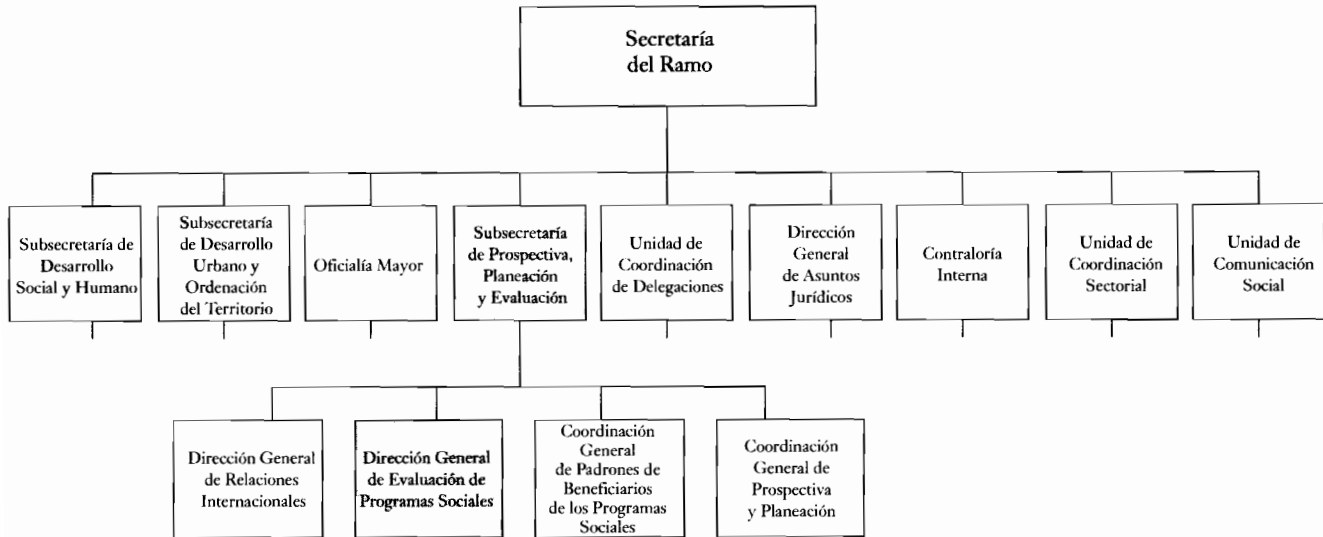
### Planeación y organización de la evaluación en Sedesol

La evaluación es organizada en la Sedesol fundamentalmente a través de la Subsecretaría de Prospectiva, Planeación y Evaluación, aunque también se desarrollan algunas actividades en la Subsecretaría de Desarrollo Urbano y Ordenación del Territorio (véase organigrama 3).

En 2003, la primera subsecretaría mencionada se componía de la Dirección General de Evaluación, la Coordinación General de Patrones de Beneficiarios de los Programas Sociales, la Coordinación General de Prospectiva y Análisis y la Dirección General de Relaciones Internacionales. Las coordinaciones si bien estaban ubicadas al mismo nivel jerárquico, tenían mayor peso político y acordaban directamente con la secretaria, pero se pensaba que pronto pasarían a ser también direcciones generales. Los aspectos financieros y presupuestales eran atendidos por la

ORGANIGRAMA 3

SECRETARÍA DE DESARROLLO SOCIAL  
(Parcial)





Dirección General de Eficiencia de Programas de Desarrollo Social, que también se pensaba modificar.

La Subsecretaría de Desarrollo Urbano y Ordenación del Territorio contaba con la Coordinación General de Microrregiones y la Dirección General Adjunta de Análisis Territorial. Ambas son responsables de la estrategia Microrregiones, que ha realizado un diagnóstico basal en materia de infraestructura municipal (con base en información cualitativa y entrevistas de percepción, cuyos datos se cruzaron con los disponibles en fuentes secundarias) y planea evaluar próximamente el impacto global de la misma.

La meta de la Dirección General de Evaluación es crear, con apoyo del BM,<sup>56</sup> un sistema integrado de monitoreo y evaluación (M&E), con base en un modelo de administración basada en resultados (ABR). Éste prevé contribuir a la creación de un sistema de evaluación multianual y multitemático, así como al fortalecimiento de la capacidad nacional en monitoreo y evaluación. Dicha dirección también es responsable de dirigir la participación de la Sedesol en el diseño del Sistema Nacional de Indicadores y de normar, coordinar, supervisar y difundir los resultados de las evaluaciones realizadas a los programas sociales por instituciones académicas y especializadas. Entre sus objetivos se encuentran:

- establecer los lineamientos y criterios para la evaluación de los programas sociales;
- diseñar el conjunto de indicadores de evaluación de impacto,
- diseñar y establecer mecanismos públicos de evaluación periódica de los programas sociales;
- normar, coordinar y supervisar las evaluaciones externas de los programas sociales;
- realizar el seguimiento de la ejecución y la evaluación de impacto socioeconómico de los programas sociales;
- evaluar el impacto socioeconómico de los programas sociales; y
- establecer y operar un sistema de información y evaluación de los programas sociales.

Su responsable ha manifestado su diagnóstico de la situación en los siguientes términos “No sabemos a ciencia cierta el impacto real de estas acciones (de apoyo a la población desfavorecida) sobre la población

<sup>56</sup> En marzo de 2004 recibió un donativo del BM por 425,000 dólares en el marco del Convenio de Cooperación Técnica para su diseño (comunicado 048, en [www.sedesol.gob.mx](http://www.sedesol.gob.mx)).

(beneficiada)...ni mucho menos tenemos un ordenamiento por eficiencia en el uso de los recursos. No hay todavía una práctica de evaluación generalizada” (Hernández, 2003).

El sistema de evaluación se justifica por su contribución a: *a)* el rediseño y la implementación de programas; *b)* el uso eficiente de los fondos públicos; *c)* el debate de políticas sociales mediante la agregación de elementos técnicos y objetivos; y *d)* la transparencia y la rendición de cuentas (Hernández, 2005).

El sistema presentado como propuesta tiene por objetivos:

- evaluar el diseño y los procesos de los programas;
- evaluar su impacto;
- evaluar la política social como un todo;
- fomentar en México la formación de capital humano en temas de evaluación; e
- institucionalizar el sistema de evaluación.

Tratándose de proyectos productivos, se buscaría que la evaluación ayude a la asignación eficiente de los recursos; y en el área de política social, que la medición de eficiencia e impacto retroalimente la pertinencia y el diseño de los programas.

El sistema se caracteriza por: *a)* poner énfasis en la medición de resultados; *b)* utilizar indicadores cuantitativos pero sin desechar las evaluaciones cualitativas; *c)* medir impacto en el largo plazo y monitorear diariamente la operación; *d)* ser realizada por buenos evaluadores externos; y *e)* ofrecer evaluaciones objetivas y útiles, con la cooperación y participación de los agentes del programa (Hernández, 2005).

## Papel del Congreso

La Cámara de Diputados cuenta con atribuciones en materia de evaluación definidas por las siguientes normas jurídicas:

- la Constitución Política de los Estados Unidos Mexicanos;
- la Ley Orgánica del Congreso General de los Estados Unidos Mexicanos;
- la Ley de Planeación;
- la Ley de Fiscalización Superior de la Federación;

- el Reglamento para el Gobierno Interior del Congreso General de los Estados Unidos Mexicanos; y
- los acuerdos parlamentarios en la materia.

La primera establece, en sus artículos 69, 79 y 93, la facultad del Congreso para expedir la ley de regulación de la ASF y normar la gestión, control y evaluación de los poderes de la Unión y los entes públicos, fiscalizar el cumplimiento de los objetivos de los programas federales a través de los informes que se rindan en los términos dispuestos por la ley, fiscalizar los recursos federales ejercidos por las entidades federativas y municipios, citar a comparecer a funcionarios (incluidos los secretarios de Estado), integrar comisiones para investigar el funcionamiento de los organismos descentralizados y empresas de participación estatal mayoritaria, etcétera.

La segunda norma citada, en los artículos 7, 39 numeral 3 y 45 numeral 1, faculta a la Cámara de Diputados a constituir comisiones ordinarias de información y control evaluatorio y de revisión de aspectos presupuestales en el marco de la cuenta pública, vinculadas al cumplimiento de los programas gubernamentales, entre otras. Sus artículos 86 y 90 fracción IX establecen la competencia del Senado para analizar y dictaminar las iniciativas de ley y de decreto relacionadas con los aspectos inherentes al desarrollo social.

La Ley de Planeación (art. 5) establece la obligación del Presidente de la República de enviar el Plan al Congreso de la Unión para su examen y opinión; este último, en el ejercicio de sus atribuciones constitucionales y legales, formulará las observaciones que estime pertinentes durante la ejecución, revisión y adecuaciones al mismo.

Por su parte, la Ley de Fiscalización Superior de la Federación norma las responsabilidades a cargo de la ASF.

El Reglamento para el Gobierno Interior del Congreso establece su organización y funcionamiento en materia de: instalación de las cámaras; presidencia, vicepresidencia, secretarios y prosecretarios; sesiones, iniciativas de leyes, comisiones, discusiones, revisiones de proyectos de ley, votaciones, fórmula de expedición de leyes, Comisión Permanente, etcétera.

Por su parte, el Acuerdo Parlamentario relativo a la Organización y Reuniones de las Comisiones y Comités de la Cámara de Diputados (art. 12, inciso *b*), define entre las funciones a cargo de comisiones, las relativas a evaluar periódicamente el PND y los programas respectivos, así como realizar audiencias, consultas, visitas, entrevistas, convocatorias a particulares y comparecencias de servidores públicos, autoridades de organismos descentralizados y empresas de participación mayoritaria, etcétera.

Las normas anteriores facultan a la citada Cámara a evaluar, cuantitativa y cualitativamente, la actuación del Ejecutivo, mediante mecanismos de evaluación, control y fiscalización, que se desarrollan en paralelo a su responsabilidad exclusiva para aprobar el Presupuesto de Egresos de la Federación (PEF); sin embargo, no definen sus atribuciones en materia de rectificación de estrategias o políticas del plan y de los programas sectoriales.

Como hemos visto, el Poder Legislativo tiene, entre otras funciones y dentro de un marco democrático, la de servir de contrapeso del Poder Ejecutivo, cuestionando y generando información alternativa a la que éste aporta en forma unilateral, ya que los informes anuales presentados por el Ejecutivo a la Cámara, ofrecen una imagen positiva de los avances realizados, al resaltar cuantitativamente las acciones implementadas en el marco de la planeación, sin referir su impacto real en la solución de los problemas. Se trata de información parcial, por lo que no permite una plena evaluación de sus planes y programas.

Prácticamente, no fue posible ejercer esta función durante el siglo pasado debido a la composición de partido único del Congreso, mismo al que también pertenecía el Presidente de la República, y a su consecuente dinámica de sumisión a la voluntad de este último. A partir de 1997, y más aún de 2000, esta situación comenzó a cambiar, al lograr cierto nivel de independencia de opinión y decisión, fruto de una composición más plural.

En este contexto, una de las tareas, constitucionalmente asignada y no desempeñada durante décadas, la constituye la evaluación del estado que guarda la administración pública federal y los resultados alcanzados por las políticas públicas. Para cumplirla como actividad permanente de control del cumplimiento de los compromisos, resultados e impactos, se requiere reactivar dos facultades básicas: los mecanismos formales de evaluación, control y fiscalización del ejercicio del presupuesto (revisión del gasto público, cumplimiento de metas anuales, etcétera) y de la ejecución de planes, programas y proyectos (logro de objetivos de los programas, eficiencia en el uso de los recursos, etcétera) por parte de la Cámara de Diputados y efectuar las tareas de control evaluatorio en las comisiones de la Cámara anterior.

Adicionalmente, es necesario innovar en los mecanismos de rendición de cuentas a la ciudadanía, ya que la glosa ha sido más un elemento de refutación partidista que de evaluación real de resultados, y tiene un impacto muy bajo en la opinión pública.

La normatividad actual permite al Congreso analizar los impactos de los programas del Ejecutivo del año fiscal anterior antes de realizar las asignaciones de presupuesto para el siguiente, facilita el análisis de la cuenta pública

y la determinación de auditorías a realizar. Resta investigar cuál es el trabajo de análisis de los informes de evaluaciones externas, y la realización directa de otras que efectivamente desarrolla, su nivel de utilización para reorientar las políticas y programas públicos, su práctica en materia de rendición de cuentas, y, en general, si todo esto está sirviendo para reactivar un debate político, abierto y participativo.

Las acciones de evaluación periódica del PND y de los programas sectoriales las desempeñan las comisiones ordinarias de la Cámara de Diputados, que utilizan instrumentos como la mencionada glosa, los informes de trabajo y las comparecencias de funcionarios; las de control financiero las realiza un órgano autónomo adscrito a la Cámara; y las de fiscalización de la cuenta pública quedan a cargo del órgano técnico de la Cámara, la ASF.

Esta Cámara contó con 40 comisiones ordinarias en la LVIII Legislatura, de las cuales las siguientes se vinculan a nuestro objeto de estudio: Presupuesto y Cuenta Pública, Desarrollo Social, y Participación Ciudadana, además de la ya mencionada ASF ([www.diputados.gob.mx](http://www.diputados.gob.mx)). También existe una Comisión de Desarrollo Social en el Senado.

La Comisión de Presupuesto y Cuenta Pública (CPCP) tiene entre sus objetivos, analizar la evolución del gasto público federal y de los temas de coyuntura que pueden tener impacto en el ejercicio presupuestal; así como estudiar el efecto en la población de los principales programas de gasto y de las alternativas para garantizar el cumplimiento de sus objetivos.

Para lograr lo anterior debe fortalecer sus relaciones de trabajo con las diversas comisiones de la Cámara y la ASF, así como recibir los informes trimestrales sobre la ejecución del presupuesto y las evaluaciones externas realizadas.

Mediante una entrevista a uno de sus miembros fue posible conocer las limitadas actividades que realizaron en torno de la evaluación de programas sociales, en su gestión fiscal 2002, debido a la falta de personal: entrega de copias de los informes de las evaluaciones externas a otras comisiones, solicitud de su análisis a diversos despachos, información de los resultados a los demás miembros de la Comisión y preparación de un resumen para entregar al presidente de la Cámara.

La Comisión de Desarrollo Social de la Cámara de Diputados se integró en septiembre de 2000 por 30 diputados, pertenecientes a distintos partidos. Contó con un presidente y cuatro secretarios. Tuvo a su cargo tareas de dictamen legislativo, de información, y de control evaluatorio, en materia de desarrollo social. Tiene también facultades para opinar sobre la propuesta

de reglas de operación que la Sedesol elabora, antes de que éstas sean enviadas para su autorización por la SHCP.

Sus integrantes se organizaron en diversas subcomisiones, entre las que se encuentra la encargada de la evaluación y seguimiento de programas. Dicha subcomisión tiene como objetivos establecer los mecanismos que permitan a la Comisión evaluar y dar seguimiento a la operación de los programas de la Sedesol, así como a la aplicación de recursos a dichos programas, aprobados en el decreto de PEF.

Su programa tentativo de trabajo, durante la LVIII Legislatura, incluyó básicamente seis actividades:<sup>57</sup>

- Dar su opinión a la Sedesol sobre los indicadores de evaluación y de gestión de los programas dirigidos al combate a la pobreza, participar en su diseño y solicitar a la Sedesol le autorice a dar su punto de vista sobre las reglas de operación.
- Opinar sobre el PND en materia de su competencia y formular observaciones a su ejecución.
- Realizar el control programático de los recursos del Ramo 20, mediante:
- Opinar y participar en la glosa del Informe Presidencial en materia de desarrollo social.
- Emitir opinión sobre la cuenta pública en materia de gestión financiera, respeto de los criterios presupuestales y cumplimiento de objetivos de los programas.
- Opinar sobre el proyecto de PEF ante la CPCP, en lo relativo al Ramo 20.

La Comisión de Participación Ciudadana fue creada el día 4 e instalada el 31 de octubre de 2000. Su principal objetivo es promover una cultura de participación ciudadana en el proceso de diseño, ejecución, seguimiento y evaluación de las políticas públicas, especialmente en las materias de estado de derecho y bienestar social. Dicha cultura se basa en principios y valores de legalidad, corresponsabilidad, inclusión, sustentabilidad, democracia, solidaridad, tolerancia, diversidad y convivencia pacífica.

Para lograr lo anterior, la comisión se propuso crear un marco jurídico que regule y promueva esta nueva cultura, proponer e impulsar iniciativas de ley, decretos o dictámenes que coadyuven a su extensión y profundización, propiciar el diálogo y la comunicación entre el Poder Legislativo

<sup>57</sup> No fue posible entrevistar a ninguno de los diputados que formaban parte de la misma en 2003, cuando fueron entregados los informes de evaluación externa a la gestión fiscal 2002, pero no parece que hayan cumplido ningún papel importante en su revisión, según lo manifestó el ex miembro de la CPCP.

y la sociedad y realizar estudios e investigaciones sobre diferentes instrumentos de participación ciudadana (referéndum, plebiscito, iniciativa popular, revocación de mandato, etcétera). Se incluiría la sistematización de la legislación existente en la materia en los ámbitos federal, estatal y municipal; la contribución al fortalecimiento de los eventos y foros de consulta regionales y nacionales, especialmente con expertos, académicos, líderes sociales y grupos de base; así como el logro de un acceso fácil a la información sobre políticas y programas gubernamentales de la sociedad civil organizada.

Como puede verse, su esfuerzo constituye una forma loable de empezar a introducir el tema pero, no aparece ningún mecanismo que informe a la población de las evaluaciones realizadas ni mucho menos que organice su participación en la gestión y evaluación de los programas gubernamentales.

Como ya se comentó, la Auditoría Superior de la Federación es el órgano técnico responsable de revisar y fiscalizar la cuenta pública, respaldado por la Ley de Fiscalización Superior de la Federación, aprobada en el año 2000. Tiene entre sus facultades la de evaluar la gestión; verificar la correcta comprobación de las erogaciones efectuadas por las entidades y empresas que reciben presupuesto público, independientemente de cualquier otra forma de control o fiscalización interna de los poderes de la Unión y de los entes públicos federales; emitir las recomendaciones que considere procedentes para fomentar la obtención de resultados y la satisfacción de las necesidades de la población; e informar con oportunidad a la ciudadanía de la fiabilidad de la rendición de cuentas, la consecución de los objetivos de los programas de gobierno, así como las irregularidades y faltas de transparencia o probidad en la gestión pública.

Su estructura orgánica se compone de:

- un auditor superior de la Federación, designado por un periodo de ocho años por el voto de las dos terceras partes de los miembros presentes de la Cámara de Diputados; y
- tres auditores especiales de cumplimiento financiero, de desempeño y de planeación e información, encargados de planear las actividades relacionadas con la revisión de la cuenta pública; ejecutarlas; designar inspectores, visitadores o auditores; instruir los procedimientos; y promover ante las actividades competentes el fincamiento de responsabilidades.

Los auditores especiales de desempeño comenzaron en 1999 a hacer evaluaciones de programas gubernamentales y sociales en particular: Progreso en su componente educativo el primer año; Jornaleros Agrícolas, Apoyo a las Empresas en Solidaridad y Progreso en su componente de salud en 2000; Ahorro y Subsidios Para la Vivienda, Crédito a la Palabra, Coinversión Social, Empleo Temporal y Progreso en su componente alimenticio en 2001; y Desarrollo de los Pueblos Indígenas y Microrregiones en 2002.

Las anteriores trascienden la mera revisión del uso de los recursos y se enfocan a la medición del desempeño en el cumplimiento de los objetivos de las instituciones y de los funcionarios y empleados públicos, con el objetivo de promover la transparencia y rendición de cuentas, identificándose plenamente con la Ley Federal de Transparencia y Acceso a la Información. A partir de 2001 pueden consultarse en su página web.

La Comisión Ordinaria de Desarrollo Social del Senado fue instalada el 11 de octubre del año 2000, con la participación de 15 senadores de tres grupos parlamentarios. Se propuso atender las propuestas y las iniciativas en materia de política social, así como impulsar una reforma constitucional que garantizara el pleno ejercicio de los derechos sociales y promoviera una ley marco de desarrollo social, misma que se concretaría, al finalizar el año 2003, con la Ley General de Desarrollo Social (LGDS), que institucionaliza el tema como política de Estado.

De esa forma se trató de fortalecer la interlocución e incidencia del Poder Legislativo en el diseño y evaluación de las políticas públicas de desarrollo social y contribuir a la construcción de consensos entre los actores políticos, sociales y económicos, como forma de alcanzar una política de desarrollo integral más acorde con las necesidades del país; así como cuidar los compromisos y acuerdos signados por el gobierno mexicano con la cooperación internacional en la materia.

## Regulación de la contratación de evaluaciones externas

### *Disposiciones generales*

Aludimos en este apartado a las disposiciones constitucionales, leyes, decretos y acuerdos que regulan la realización de las evaluaciones externas.

El decreto de PEF del año 2000 señalaba, en su artículo 73, que las dependencias responsables de coordinar programas sujetos a reglas de operación que no contaran con un mecanismo de evaluación externa debían integrar un Consejo Técnico de Evaluación y Seguimiento que incluyera instituciones



académicas. En consecuencia, hasta ese año la evaluación externa se consideraba deseable más no obligatoria.<sup>58</sup>

A partir del ejercicio fiscal 2001, el Congreso comenzó a exigir la realización de evaluaciones externas a los programas que reciben recursos fiscales y están sujetos a reglas de operación, cuya lista se incluye en el PEF. Las mismas se fundamentaron en los artículos: 134 de la Constitución Política de los Estados Unidos Mexicanos; 26 fracción II, 40, 41, 42 y 43 de la Ley de Adquisiciones, Arrendamientos y Servicios del Sector Público; y 70 del PEF. Para el caso del ejercicio fiscal 2002, ésta aparece en el artículo 63, fracción V, inciso *b* del PEF.

Las reglas de operación son diseñadas por la secretaría responsable de los programas y autorizadas por la SHCP, para garantizar la aplicación eficiente, eficaz, equitativa y transparente de los recursos públicos (artículos del PEF: 74 en el año 2000, 71 en el 2001 y 64 en el 2002).

La Constitución, por su parte, establece la obligación del gobierno federal de administrar con eficiencia, eficacia y honradez los recursos económicos de que disponga para la satisfacción de objetivos. Para lograrlo exige la realización de licitaciones públicas, mediante convocatoria también pública, en los casos de prestación de servicios como son las evaluaciones externas, para asegurar las mejores condiciones de precio, calidad, financiamiento, oportunidad y demás circunstancias pertinentes.

En ese sentido, la Ley de Adquisiciones, Arrendamientos y Servicios del Sector Público señala en su artículo 26 que las dependencias y entidades, bajo su responsabilidad, podrán contratar adquisiciones, arrendamientos y servicios a organizaciones ajenas a la administración pública mediante los procedimientos de contratación: licitación pública, invitación a cuando menos tres personas, o adjudicación directa.

Agrega que, en los procedimientos de contratación, deberán establecerse los mismos requisitos y condiciones para todos los participantes, especialmente por lo que se refiere a tiempo y lugar de entrega, forma y tiempo de pago, penas convencionales, anticipos y garantías; debiendo las dependencias y entidades proporcionar a todos los interesados igual acceso a la información relacionada con dichos procedimientos, a fin de evitar favorecer a algún participante.

Su artículo 40, a su vez, establece que, en los supuestos que prevé el artículo 41 de dicha ley, las dependencias y entidades, bajo su responsabilidad, podrán optar por no llevar a cabo el procedimiento de licitación pública

<sup>58</sup> Sin embargo, el BM (2004a) señala que fue a partir de ese año que el PEF exigió la realización de evaluaciones externas.

y celebrar contratos a través de los procedimientos de invitación a cuando menos tres personas o de adjudicación directa.

La selección del procedimiento que realicen las dependencias y entidades deberá fundarse y motivarse, según las circunstancias que concurren en cada caso, en criterios de economía, eficacia, eficiencia, imparcialidad y honradez que aseguren las mejores condiciones para el Estado. La acreditación de los criterios mencionados y la justificación de las razones para el ejercicio de la opción, deberá constar por escrito y ser firmado por el titular del área usuaria o requirente de los bienes o servicios.

En cualquier supuesto se invitará a personas que cuenten con capacidad de respuesta inmediata, así como con los recursos técnicos, financieros y demás que sean necesarios, y cuyas actividades comerciales o profesionales estén relacionadas con los bienes o servicios objeto del contrato a celebrarse.

En estos casos, el titular del área responsable de la contratación, a más tardar el último día hábil de cada mes, enviará al órgano interno de control en la dependencia o entidad de que se trate, un informe relativo a los contratos formalizados durante el mes calendario inmediato anterior, acompañando copia del escrito aludido en este artículo y de un dictamen en el que se hará constar el análisis de la o las propuestas y las razones para la adjudicación del contrato.

El artículo antes citado establece 18 situaciones en las que es posible evitar la licitación pública, de las cuales nos interesan los numerales tres, siete y 10, que pueden resultar aplicables en nuestro caso:

- existan circunstancias que puedan provocar pérdidas o costos adicionales importantes, debidamente justificados;
- se realicen dos licitaciones públicas que hayan sido declaradas desiertas; y
- se trate de servicios de consultorías, asesorías, estudios e investigaciones cuya difusión pudiera afectar al interés público o comprometer información de naturaleza confidencial para el gobierno federal.

El artículo 42 agrega a los casos anteriores una nueva situación: cuando el importe de cada operación no exceda los montos máximos que al efecto se establecerán en el PEF, no pudiendo exceder del 20 por ciento del presupuesto total de adquisiciones de la dependencia o entidad, salvo casos excepcionales. Dichos montos fueron fijados en el artículo 62 del PEF 2000, 58 del PEF 2001 y 51 del PEF 2002. Para el último año, suponiendo que el volumen anual de adquisición presupuestado fuera mayor de 1,000 millones

de pesos, las dependencias y entidades podían contratar hasta 300,000 pesos por adjudicación directa y 2'300,000 pesos por invitación a cuando menos tres personas.

El artículo 43 señala que el procedimiento de invitación a cuando menos tres personas se sujetará a lo siguiente:

- I. El acto de presentación y apertura de proposiciones se llevará a cabo en dos etapas, para lo cual la apertura de los sobres podrá hacerse sin la presencia de los correspondientes licitantes, pero invariablemente se invitará a un representante del órgano interno de control en la dependencia o entidad;
- II. Para llevar a cabo la adjudicación correspondiente, se deberá contar con un mínimo de tres propuestas susceptibles de analizarse técnicamente;
- III. En las invitaciones se indicarán, como mínimo, la cantidad y descripción de los bienes o servicios requeridos, plazo y lugar de entrega, así como condiciones de pago;
- IV. Los plazos para la presentación de las proposiciones se fijarán para cada operación atendiendo al tipo de bienes o servicios requeridos, así como a la complejidad para elaborar la propuesta;
- V. El carácter nacional o internacional en los términos del artículo 28 de esta ley, y
- VI. A las demás disposiciones de esta ley que resulten aplicables.

Si bien los términos de referencia revisados se basan en los artículos mencionados, esta ley prevé su no obligatoriedad (artículo 1o.) en los casos en que la contratación se realiza entre entidades y/o dependencias de la propia administración pública.

La Ley de Adquisiciones cuenta con un reglamento que especifica más detalladamente estos procedimientos y entró en vigor el 19 de septiembre de 2001.

El PEF establece también que la evaluación externa se ha de entregar a más tardar el 15 de octubre a la Comisión de Presupuesto y Cuenta Pública de la Cámara para que “los resultados sean considerados en el proceso de análisis y aprobación del Presupuesto de Egresos de la Federación para el siguiente Ejercicio Fiscal” (*Diario Oficial de la Federación*, 1o. de enero de 2002: 71).<sup>59</sup>

<sup>59</sup> Este plazo será modificado al entrar en vigor la LGDS: abril a marzo de cada año, en lugar de enero a diciembre.

Señala además que las evaluaciones deberán hacerse con cargo a los presupuestos de los programas correspondientes, por instituciones académicas y de investigación u organismos especializados, de carácter nacional o internacional,<sup>60</sup> que cuenten con reconocimiento y experiencia en las respectivas materias de los programas.

Los requisitos mínimos que deberán cumplir las organizaciones interesadas en realizar la evaluación de los programas gubernamentales sujetos a reglas de operación, a que se refiere el artículo 64 del decreto antes mencionado, se especifican en el artículo tercero del Acuerdo, firmado por la SHCP y la entonces Secodam (en adelante el Acuerdo), publicado en el *Diario Oficial de la Federación* (30 de abril de 2002, segunda sección: 4-7); mismo que continuó vigente en 2003, según se establece en el artículo 8o. transitorio del PEF correspondiente:

- I. Acreditar su constitución legal en los términos de la legislación aplicable, y en su caso, acreditar además, tener reconocimiento oficial expedido por autoridad competente;
- II. Acreditar experiencia en la evaluación de programas gubernamentales, en México o en el extranjero, en los términos del inciso b, numeral 2 de la fracción siguiente, y
- III. Presentar una propuesta de trabajo ejecutiva, que contenga, entre otros, los siguientes aspectos:
  - a) La metodología de evaluación específica a implementar, la estructura temática del informe a elaborar con una breve descripción de su contenido, y la enunciación de los trabajos de gabinete, de campo y, en su caso, las encuestas que se realizarán;
  - b) La currícula del personal que realizará la evaluación del programa gubernamental de que se trate, que incluya:
    1. La manifestación por escrito de que se tiene conocimiento de las características y operación del programa objeto de evaluación, o bien, de programas similares;
    2. La acreditación de experiencia en la evaluación de programas similares, incluyendo, en su caso, los reconocimientos a nivel nacional y/o internacional sobre este tipo de trabajos;

<sup>60</sup> El PEF 2004 presenta esta parte modificada en relación al carácter de las instituciones y organismos citados, que sólo podrán ser nacionales; en cambio la Dirección General de Metodología promueve la colaboración entre evaluadores nacionales e internacionales (Hernández, 2005). La LGDS prohíbe también la participación de empresas consultoras.

3. Los responsables de evaluar los resultados del programa deberán contar con título profesional o grado académico equivalente;
- c) La descripción técnica de la infraestructura de recursos informáticos del evaluador externo que utilizará en la evaluación de los resultados del programa;
- d) Los evaluadores externos precisarán, en su caso, cuáles serán las áreas responsables de evaluar los resultados del programa;
- e) La definición de la plantilla de personal que se utilizará para evaluar los resultados del programa, la cual deberá guardar congruencia con la magnitud y características particulares del programa.

En relación al contenido que los informes elaborados por los evaluadores externos deben incluir, los mismos se especifican en el artículo 4o. del Acuerdo, que establece sus requisitos mínimos. Se trata de los siguientes:

- I. El porcentaje de cumplimiento de los objetivos generales y particulares de cada programa, destacando principalmente:
  - a) El grado de satisfacción de la población objetivo del programa evaluado;
  - b) La cobertura alcanzada por el programa en el ejercicio fiscal 2002, y su variación con respecto a la obtenida en el año anterior, por tipo de apoyo otorgado:
    - A nivel nacional;
    - Por entidad federativa;
    - Por núcleo de población o tipo de empresa;
- II. El cumplimiento de las reglas de operación que regulan el programa de que se trate. Asimismo, en caso de que se hayan realizado modificaciones a dichas reglas en los términos del artículo 63, Fracción IV, del Decreto de Presupuesto, la manifestación respecto de las circunstancias extraordinarias o los problemas operacionales que dieron motivo a las modificaciones;
- III. En términos de impacto, los beneficios económicos y sociales que resultan de la aplicación de los recursos asignados a cada programa;
- IV. El costo-beneficio de los recursos asignados al programa, en función de costos y gastos unitarios en que se incurre para operar el programa;
- V. El resultado de la verificación en campo del cumplimiento de las metas físicas y financieras del programa;
- VI. La presentación del escenario prospectivo y la opinión sobre la permanencia o conclusión del programa, y

VII. Conclusiones y recomendaciones, que destaquen particularmente los aspectos enunciados en este artículo, así como las deficiencias que, en su caso, pueda tener el programa o sus reglas de operación y las posibles medidas para subsanarlas.

Debido a las diferentes interpretaciones que del Acuerdo anterior hacen los evaluadores, en julio de 2003, la SFP elaboró una guía para la realización de evaluaciones externas que pretende interpretar y aclarar algunos de sus aspectos, presentando avances en algunos puntos y sin lograrlo plenamente en otros. Se trata de un documento en que se revisa el alcance de la evaluación, su nivel de representatividad y el contenido del informe:

- Aporta una definición operativa de cobertura, que el Acuerdo no aclaraba, en tanto “la proporción que existe entre población objetivo que recibe el programa entre la población objetivo total” (p. 6).
- Precisa el concepto de focalización como “la concentración de los recursos de éste (el programa) de acuerdo a las características de la población objetivo del programa” (p. 6).
- Exige un análisis detallado del proceso de muestreo, descuidado en el Acuerdo (p. 4).
- Reconoce que por la naturaleza de algunos programas el impacto de los beneficios económicos y/o sociales no se puede medir a corto plazo.
- Destaca la referencia a que “La información deberá ser representativa por programa, entendiéndose por programa...” (p. 4), pero sin aclarar el concepto relevante, el de representatividad, y su factibilidad con base en los recursos disponibles.
- No alude a los casos de evaluación más complejos, referidos a la medición de servicios intangibles como los vinculados a la cultura, sino que se ejemplifica con los más sencillos.
- No parece percibir la imposibilidad de que el evaluador verifique el número total de obras y beneficiarios.
- Parece desconocer que la mayoría de los programas carecen de un padrón de beneficiarios.
- Confunde el método de costo-eficacia con costo-eficiencia y explica el análisis costo-beneficio en su versión *a priori*, muy diferente de la medición de resultados e impacto.

En mayo de 2004, la SFP elaboró una propuesta de nuevo acuerdo con la SHCP que pretende sustituir el de abril de 2002. Presenta importantes avances

como incluir la evaluación del diseño del programa o espaciar las mediciones de impacto. Incluye también algunos aspectos no compartidos en este trabajo como la eliminación de la verificación en campo del logro de metas, y no refiere otros considerados importantes como los recursos financieros y de información que deberían estar disponibles para que el evaluador pueda desarrollar su trabajo, así como muestra cierta imprecisión en la terminología empleada.

### *Bases y términos de referencia*

Las bases estipulan la forma en que se desarrollará el proceso de presentación de propuestas técnicas y económicas, selección y contratación de evaluadores, condiciones de pago, sanciones, etcétera.

En ellas se establecen los criterios que se aplicarán para seleccionar las propuestas, entre los que destacan: sólo se evalúan las propuestas que cumplan con todos los requisitos solicitados, se seleccionan las que garanticen la realización del trabajo y, dentro de ellas, la que proponga el precio más bajo. En caso de empate se determina por sorteo.

Las bases informan también sobre las reuniones de aclaración de las mismas, apertura de ofertas técnicas, lectura del dictamen sobre las propuestas anteriores, apertura de las ofertas económicas y fallo final, en que los licitantes pueden estar presentes.

Los términos de referencia proponen una definición de la evaluación, que sirve de guía al trabajo a realizar. Algunas organizaciones ofrecen definiciones más detalladas; por ejemplo, para el INI, la evaluación es un mecanismo de aprendizaje institucional que reúne evidencias tangibles, documentales y de campo, de los logros del programa y sus proyectos, a fin de contar con información oportuna, veraz y confiable para: *a)* el análisis y aprobación del presupuesto 2003; *b)* la detección de desviaciones e impactos, para establecer mejoras e innovaciones; y *c)* la determinación de los criterios para su continuidad, cancelación o rediseño y, en su caso, las estrategias necesarias para su consolidación.<sup>61</sup>

En el caso de Diconsa, se espera la aplicación de metodologías y procedimientos que analicen su eficiencia, eficacia, costos, beneficios económicos y sociales, calidad, distribución equitativa y transparente de los recursos asig-

<sup>61</sup> Véase términos de referencia de cualquiera de los programas del INI (2002), en particular, Programa de Planeación, pp. 11 y 16. Para 2004 esta definición se precisa en los términos de referencia de iniciativa ciudadana 3x1 (p. 19) como una evaluación global en términos de Baker (2000), incluyendo la supervisión, evaluación de los procesos, evaluación de costos-beneficios y evaluación de impacto.

nados, grado de satisfacción de los beneficiarios, apego a las reglas de operación, e impacto social con el propósito de mejorar sus acciones y garantizar la confiabilidad del cumplimiento de los objetivos del programa.<sup>62</sup>

Finalmente, los citados términos establecen los métodos que deberán emplearse en el trabajo. Nuevamente para el INI, éstos serán cuantitativos y cualitativos, siendo los resultados estadísticos los más relevantes, e incorporando técnicas antropológicas y sociales como observación directa, entrevistas y cuestionarios, con base en el diseño de una muestra representativa aportada por Sedesol, además del análisis documental realizado en gabinete. Esta información se maneja mediante el diseño de indicadores.<sup>63</sup>

En cuanto al procedimiento, la entidad contratante es la encargada de recibir los informes y entregarlos a la SHCP, la Secodam-SFP y la Sedesol. Esta última lo entrega, en su carácter de coordinadora del sector, a la Cámara de Diputados.

## Ley Federal de Transparencia y Acceso a la Información Pública Gubernamental

A pesar de toda la importancia que el PND y el PNDS otorgan a la participación social y ciudadana, revisada en el séptimo capítulo, el papel de la población no organizada en la evaluación de políticas y programas sociales no ha pasado de ser “objeto” de encuestas a “sujetos” activos dentro de los procesos participativos y de aprendizaje, postulados tanto por los organismos internacionales como por los gobiernos nacionales.

Frente a esta situación, las democracias modernas han buscado mecanismos alternativos de participación ciudadana directa. Pero, para que la ciudadanía pueda evaluar los programas del gobierno requiere tener acceso a la información a través de procesos de rendición de cuentas. Ésta se define como “...la obligación que tienen quienes ejercen el poder público de responsabilizarse de su labor, de someterse a evaluaciones de desempeño y de dar a conocer los resultados de esa evaluación” (López: 2002: 5) y tiene por objeto evitar abusos de los poderes en beneficio personal de quienes los ejercen.

Dicha rendición de cuentas se ha empezado a organizar a partir de junio de 2002, en que fue aprobada la Ley Federal de Transparencia y Acceso a la Información Pública Gubernamental. Esta ley establece que “...toda información gubernamental a que se refiere esta Ley es pública y

<sup>62</sup> Consultar términos de referencia para el Estudio de Evaluación Externa del Programa de Abasto Rural, 2002, p. 11 y Procedimiento de Invitación, p. 2.

<sup>63</sup> Posteriormente la Sedesol informó que se referían a “indicadores” como sinónimo de temas a incluir.



los particulares tendrán acceso a la misma en los términos que ésta señala” ([www.secodam.gob.mx/leyes](http://www.secodam.gob.mx/leyes): artículo 21) y prevé un plazo de un año para que todas las entidades obligadas (tanto organismos de la administración central como los de carácter autónomo) pongan su información al alcance del público.

A esta medida la ley la llama “obligaciones de transparencia” y la información a brindar queda especificada de la siguiente manera: “Con excepción de la información reservada o confidencial prevista en esta Ley, los sujetos obligados deberán poner a disposición del público y actualizar, entre otra, la siguiente información:

- su estructura orgánica,
- las facultades de cada unidad administrativa,
- el directorio de servidores públicos, desde el nivel de jefe de departamento o sus equivalentes,
- la remuneración mensual por puesto, incluso el sistema de compensación, según lo establezcan las disposiciones correspondientes,
- las metas y objetivos de las unidades administrativas de conformidad con sus programas operativos,
- los servicios que ofrecen” ([www.secodam.gob.mx/leyes](http://www.secodam.gob.mx/leyes): artículo 7).

Además de los servicios, los organismos incluidos en la ley deben informar los trámites que llevan a cabo, el presupuesto y su operación, los órganos de control, sus datos para ponerse en contacto, su fundamento jurídico, los informes que están obligados a presentar, etcétera.

Sin embargo, entre las obligaciones señaladas por la Ley de Transparencia no se incluye información sobre los resultados de evaluaciones realizadas, por lo que dicha información debe ser solicitada al Instituto Federal de Acceso a la Información (IFAI).<sup>64</sup> Esto a pesar de que el propio BM (2004b) expresa la importancia de la difusión de los resultados de la evaluación para facilitar la comunicación entre los beneficiarios, la sociedad civil y el gobierno.

El IFAI está obligado a responder a toda solicitud de información a través de su unidad de enlace. Asimismo, desarrolla diversas acciones de información acerca de sus objetivos, su naturaleza y sus acciones, a través de sus programas de comunicación social (que son masivos) y de sus programas de vinculación con la sociedad (con organizaciones civiles) o de manera perso-

<sup>64</sup>Esto cambiará con la entrada en vigencia de la LGDS que exige la publicación de los resultados de las evaluaciones externas realizadas.

nalizada, al ciudadano que así lo requiere, y verifica que todo lo previsto por la Ley de Transparencia se cumpla con claridad y veracidad.

Puesto que el artículo 7o. de la ley dice que son públicas “la información sobre el presupuesto asignado, así como los informes de su ejecución...y los informes que generen los sujetos obligados”, un funcionario entrevistado en el IFAI concluía que, si los informes de evaluación externa analizan el desempeño y la calidad de las funciones de la institución, y sus costos son financiados con recursos públicos, entonces pueden ser conocidos por el público.<sup>65</sup>

## Ley General de Desarrollo Social

La LGDS, aprobada a finales de 2003, establece la obligación legal de realizar estas evaluaciones, por lo que ya no va a ser necesario que el PEF lo reitere cada año. Además, prevé en sus artículos 72 y 81 la creación del Consejo Nacional de Evaluación de la Política de Desarrollo Social (CNE) como organismo público descentralizado, encargado de revisar periódicamente el cumplimiento de objetivos de programas y acciones, por sí mismo o a través de uno o varios organismos, independientes del ejecutor del programa. Dichos organismos de evaluación pueden ser instituciones de educación superior, de investigación científica u organizaciones no lucrativas (art. 73), y será responsabilidad del Consejo hacer su selección mediante procesos de licitación.

El CNE se integra por el titular de la Sedesol o la persona que éste designe, seis investigadores académicos y un secretario ejecutivo designado por el Ejecutivo federal. Los investigadores, que duran cuatro años en el cargo pudiendo reelegirse la mitad de ellos, deben ser o haber sido miembros del Sistema Nacional de Investigadores, con amplia experiencia en la materia, ser colaboradores de instituciones de educación superior y de investigación, inscritas en el Padrón de Excelencia del Consejo Nacional de Ciencia y Tecnología. Son designados por la Comisión Nacional de Desarrollo Social, también creada por esta ley (arts. 82 y 83).

La citada ley agrega que se buscará medir cobertura, calidad e impacto y se emplearán indicadores de resultados, gestión y servicios. Los primeros de estos reflejarán el cumplimiento de objetivos sociales; los restantes, los procedimientos y calidad de los servicios (arts. 74 a 76). Los indicadores deberán ser aprobados por el CNE, después de solicitar su opinión y recomendaciones a la SHCP y la Cámara de Diputados (art. 77). La evaluación será anual y cubrirá

<sup>65</sup> Éste se verificó luego en la práctica.

el periodo del 1o. de mayo al 30 de abril, o multianual (art. 78). Los resultados de las evaluaciones serán publicados en el *Diario Oficial de la Federación* y se entregarán a las comisiones de desarrollo social de las cámaras de Diputados y Senadores del Congreso de la Unión y a la Sedesol (art. 79). Con base en dichos resultados, el CNE podrá emitir sugerencias y recomendaciones al Ejecutivo federal y hacerlas del conocimiento público (art. 80).

A partir de su entrada en vigencia, las comisiones de desarrollo social de la Cámara de Diputados y del Senado también recibirán copias de los informes de evaluación externa.

Esta ley enfatiza la importancia de la participación social y ciudadana a través de instrumentos como la contraloría social.

### Comentarios finales

El capítulo permite contar con el marco normativo en que se organizan y planean los trabajos de evaluación de los programas de desarrollo social, mediante la descripción de las funciones que corresponden a las diversas instancias participantes de la administración pública federal y el Congreso, así como la forma en que se regula la contratación de las instituciones académicas y de investigación que actúan como evaluadoras. También revisó la legislación vigente para facilitar la participación social y comunitaria, que aparece como una de las estrategias privilegiadas en el PND y el PNDS.

En relación a la normatividad actual de los procesos de evaluación resaltan los recientes avances en materia de expedición de leyes como: Ley Federal de Acceso a la Información Pública Gubernamental, Ley General de Desarrollo Social y también la Ley Federal de Fomento a las Actividades Realizadas por las Organizaciones de la Sociedad Civil.

En cambio surgen como aspectos poco claros o negativos: el retraso incurrido en la expedición del reglamento de la LGDS y la creación del CNE, la falta de precisión en los términos del Acuerdo que se reflejan en los términos de referencia; la obligación establecida por el PEF en el sentido de que la evaluación debe realizarse con cargo al programa evaluado, lo que pone en riesgo la independencia del evaluador; y la falta de una reglamentación clara que establezca cuándo se puede otorgar un contrato por adjudicación directa a una entidad o dependencia de la administración pública en lugar de abrir una licitación pública y que defina límites a los precios que pueden pagarse en esa situación.

La primera impresión que deja la revisión realizada es la existencia de un amplio número de organizaciones con responsabilidades formales en materia

de evaluación; no obstante, resta analizar el grado en que éstas se cumplen en la práctica y revisar si también están vigentes otros mecanismos de carácter informal, lo que se revisará en el siguiente capítulo, dedicado al análisis del contexto real en que se desarrolla la evaluación de programas sociales. El mismo es el primero de tres capítulos dedicados a encarar la metaevaluación de programas de desarrollo social desde diferentes ángulos.

## Capítulo 10

# Organización y gestión del sistema de evaluaciones externas

### Introducción

ESTE ES EL primero de tres capítulos dedicados a la metaevaluación de los programas sociales en México, en los que se analiza el origen, las condiciones, los procesos, los resultados y los impactos de las evaluaciones aplicadas a programas sociales, así como la calidad del trabajo realizado y sus respectivos costos.

Siguiendo a Duran (2002) y Chanut (2002), presentados en el segundo capítulo, se considera relevante en una metaevaluación la revisión de los elementos del contexto que pueden dificultar o facilitar el desarrollo de las evaluaciones y contribuir u obstaculizar la buena calidad de sus resultados.

Se hace referencia tanto al contexto global en el que se inserta el sistema de evaluación, como el contexto inmediato de los procesos que su trabajo implica, compuesto por todas las instancias que planean, organizan y gestionan las evaluaciones externas. Esto se realiza con base en el diseño de indicadores temáticos de análisis o listas de verificación, presentados en nuestra propuesta del sexto capítulo.

El capítulo se apoya en información secundaria y en la aplicación de entrevistas semiestructuradas a los diversos actores-clave vinculados al proceso: autoridades de Sedesol, funcionarios de las instancias operadoras de los programas, funcionarios de otras secretarías, de institutos vinculados a la evaluación y rendición de cuentas, evaluadores externos, integrantes de organizaciones de la sociedad civil y miembros del Senado y la Cámara de Diputados.

### Contexto político y cultural

En los últimos años, el PRI dejó de ser un partido de Estado. La oposición fue poco a poco ganando gubernaturas estatales, presidencias municipales y presencia en los congresos; sin embargo, por primera vez, después de siete

décadas, el Partido Acción Nacional (PAN) ganó las elecciones a la Presidencia de la República en julio de 2000. En particular, el gobierno del Distrito Federal ha continuado en manos del PRD desde 1994 y su relación con el gobierno federal se ha vuelto mucho más ríspida en el periodo en curso.

Desde las elecciones del año 2000, el Poder Legislativo quedó integrado por miembros de los distintos partidos, sin que el PAN obtuviera la mayoría absoluta, lo que hacía prever importantes debates y negociaciones para la aprobación de iniciativas y un contrapeso relevante a las acciones del Ejecutivo.

En paralelo a esta democracia representativa, el gobierno entrante propuso procesos intensivos de participación ciudadana directa, por vía de la consulta diagnóstica y propositiva (revisados en el capítulo 3), y las OSC incrementaron sus exigencias para incidir en los procesos de políticas públicas, al mismo tiempo que los medios masivos de comunicación ganaban espacio y poder.

Los problemas para lograr la participación directa han incluido la voluntad política para generar los canales, el acceso a la información y la desigualdad de capacidades de los actores sociales para realizar sus análisis.

Es en este contexto político que el Ejecutivo saliente en el año 2000 habría propuesto la aplicación de evaluaciones a los programas gubernamentales, estrategia compartida por el BM y el BID. La situación cumplía con las condiciones, analizadas en el sexto capítulo, en que se inician procesos de evaluación: un fuerte grado de incertidumbre o un elevado potencial de conflicto (Nioche, 1982), la complejización de la acción pública y las dificultades del aparato estatal para justificarse y legitimarse (Duran y Monnier, 1992).

Complementariamente, en el año 2001 se puso en marcha el Programa Nacional para la Transparencia y el Combate a la Corrupción, que entre sus líneas estratégicas incluye: mecanismos de rendición de cuentas, de supervisión y control, y de participación ciudadana.

Con la intención de responder a la necesidad de transparencia que requiere una ciudadanía activa, en 2002 se aprobó la Ley Federal de Transparencia y Acceso a la Información a la Información Pública Gubernamental, misma que no ha impactado considerablemente la difusión de resultados de las evaluaciones realizadas.

La cultura política desarrollada durante tantos años de partido hegemónico en el poder no había generado incentivos para impulsar estrategias de evaluación de sus políticas y programas, por lo que sólo había desarrollado

algunos mecanismos de control punitivo, que se habían encargado de provocar temor y rechazo por parte de los funcionarios inspeccionados.

Ese conjunto de hábitos, conocimientos, creencias, ideas, valores, normas, etcétera que integran la cultura de la evaluación de un país enfrentaba los tres tipos de problemas señalados por Nioche (1982) para el contexto francés (revisados en el capítulo 6): sociopolíticos, administrativos y metodológicos.

Los primeros debido al débil rol del Poder Legislativo como evaluador de la acción gubernamental, el fuerte control de los medios de comunicación masivos por el Ejecutivo, el cuasimonopolio de la información económica y social por la administración pública, y el temor provocado por la duda y la posibilidad de cuestionamiento entre los altos funcionarios, acostumbrados a comportamientos informales.

Los obstáculos administrativos provocados por la debilidad de la gestión por programas; el diseño de planes, programas y presupuestos con objetivos muy vagos y ausencia de metas cuantificadas que faciliten su evaluación; la falta de recursos destinados a esta actividad, cuya aportación al bienestar social era muy discutida; y el rechazo de algunos funcionarios a una evaluación que consideran inútil.

Los problemas metodológicos surgidos como consecuencia de que todas las técnicas empleadas en evaluación enfrentan cuestionamientos que disminuyen su credibilidad, como fuera analizado en el tercer capítulo.

En este contexto político y cultural dio inicio, en 2001, la obligación impuesta por la Cámara de Diputados de realizar evaluaciones externas a los programas gubernamentales sujetos a reglas de operación.

## Tipos de evaluación planeados

Para alcanzar sus objetivos el sistema de evaluación cuenta con siete componentes:

- a) Evaluación externa de corto plazo: es la que se está realizando actualmente, a cargo de instituciones académicas y especializadas, para responder a la obligación impuesta por la Cámara de Diputados, e incluye los aspectos de focalización, gestión, apego a las reglas de operación e impacto.
- b) Evaluación externa de largo plazo: se propuso su evaluación multianual (2003-2005/2006), también desarrollada por instituciones académicas y especializadas, para medir el impacto real de los programas en la población.

- c) Evaluación interna de diseño e impacto: a desarrollar por la Sedesol en los programas de Microrregiones, Jornaleros Agrícolas y Tu casa (denominado Ahorro y Subsidios para la Vivienda durante 2002), para complementar la evaluación externa.
- d) Evaluación interna de existencia de obras y sustentabilidad: verificación por muestreo, por instancia no especificada, de pertinencia, acciones, obras, avance de compromisos, sustentabilidad de las obras después de uno o dos años de construidas y calidad del gasto social.
- e) Evaluación de la política social: mediante la cooperación técnica del BID y el apoyo de alguna universidad de excelencia como Harvard, se planteó que se trataría de analizar las duplicidades, complementariedades y sinergias de diversos programas, la relación entre las acciones de Sedesol y los impactos observados, así como la evaluación conjunta de aquellos programas en que esto sea factible. Se sugirió al respecto una evaluación de diversas localidades con distintas combinaciones de programas con más de tres años de impacto.
- f) Evaluación cualitativa: se mencionó que se llevaría a cabo durante el año 2002, aunque tampoco se especificó la instancia responsable, mediante encuestas a la población en general y a los beneficiarios en particular para conocer las percepciones que se tienen de la política social y los programas (Oportunidades, Tu Casa, Abasto Social de Leche y Abasto Rural) de la institución; también se incluiría la percepción de la propia Sedesol.<sup>66</sup>
- g) Fomento a la formación de capital humano en temas de evaluación: se propuso capacitación interna a la dirección y externa (universidades), para mejorar la cantidad y calidad de expertos nacionales en evaluación de programas sociales, a través de cursos, seminarios, talleres, ponencias, becas del Consejo Nacional de Ciencia y Tecnología (Conacyt), etcétera.

Adicionalmente, se planteó que se trabajaría en el sentido de institucionalizar el Sistema de Evaluación, para lo cual será necesario realizar su metaevaluación, por parte de una institución externa con asistencia técnica constante (BM, BID), en el periodo de junio de 2003 hasta el 2005 y obtener la certificación ISO 9000, de manera de garantizar su sustentabilidad.

<sup>66</sup> La doctora Graciela Tervel, de la Universidad Iberoamericana, presentó resultados de un trabajo de este tipo, "Encuesta nacional sobre niveles de vida de los hogares" en el seminario de desarrollo social, realizado en Sedesol el 19 de noviembre de 2004.



Por su parte, la Dirección General Adjunta de Análisis Territorial, como ya se mencionó, busca generar impactos en la infraestructura y servicios de las microrregiones, no trabaja a nivel de los individuos, sino de su confluencia en un territorio. En consecuencia el impacto de su estrategia no es la mera suma de los impactos de los programas aplicados y requiere ser evaluada como tal, lo que aún no ha podido concretarse.

Incluso algunos programas desarrollan otras evaluaciones, también externas o internas. Es el caso del proyecto FIDA-Península de la Comisión Nacional para el Desarrollo de los Pueblos Indígenas (INI en 2002), de los programas operados por Liconsa o del programa Oportunidades, por ejemplo.

Finalmente, existen las auditorías de desempeño realizadas desde 1999 por la ASF, con un promedio de 24 trabajos por año, ocho de los cuales corresponden al sector social (Sedesol, SEP y SSA). En el periodo 1999-2001 evaluaron Progresá en su parte educativa, de salud y alimentaria. Ese último año también auditaron Crédito a la Palabra, Atención a Jornaleros Agrícolas y los programas del Fondo Nacional de Empresas en Solidaridad (Fonaes), que en ese momento dependían de Sedesol. En 2002 hicieron lo propio con Fondos Regionales Indígenas (programas a cargo del INI), Microrregiones y Desarrollo de los Pueblos y Comunidades Indígenas. Las mismas se incluyen en el informe de resultados de la cuenta pública.

En síntesis, la evaluación externa constituye un subconjunto muy importante de las evaluaciones planeadas por Sedesol y el Congreso, pero como puede verse, no la única. Se insiste en que el trabajo que se presenta a continuación sólo analiza las evaluaciones externas realizadas para cumplir la obligación establecida por la cámara de diputados, que son las más numerosas.

En términos de las clasificaciones básicas revisadas en el capítulo 2, dichas evaluaciones se clasifican como:

- a) Parcialmente formativas, ya que realizan su análisis con fines de mejoramiento de los programas pero, en la práctica, son poco utilizadas.
- b) *Ex post*, debido a que su valoración se realiza después de acabada dicha gestión.
- c) Externas, porque los evaluadores carecen de vinculación directa con los operadores de los programas, funcionarios y congresistas.
- d) De diseño, proceso, resultados e impactos: si bien no todas las evaluaciones cubren todos estos aspectos (de hecho la evaluación del diseño ni siquiera se solicita en el Acuerdo y los Términos de Referencia), el conjunto de evaluaciones sí lo hace.

Habría que agregar que se trata además de evaluaciones anuales, realizadas partir del año 2001, salvo en el de Oportunidades (antes Progresá) que las venía realizando desde 1998.

### Decisión de realizar evaluaciones externas

Como ya se ha mencionado, la decisión de evaluar los programas gubernamentales sujetos a reglas de operación fue tomada al aprobarse el PEF que se aplicaría en el año 2001. Sin embargo, el proyecto de presupuesto entregado en el año 2000 por la SHCP a la Cámara de Diputados, en vísperas del cambio en la composición partidista de los poderes Ejecutivo y Legislativo, incluyó la decisión de generalizar los procesos de evaluación de programas federales.

Considerando las presiones de tiempo con que se suele aprobar el PEF, la concentración de la atención de los legisladores en la asignación de los recursos presupuestales y el hecho de que el punto no aparece discutido en los diarios de debate de la Cámara, habría que suponer que fue el Ejecutivo saliente quien tomó a conciencia la decisión de evaluar, mientras fue el Legislativo quien formalmente la aprobó, sin estudiar a fondo sus consecuencias.

Lo anterior concuerda con la opinión de autoridades de Sedesol, para quienes se trató de una iniciativa de los operadores de los propios programas y de la SHCP, en que la última sentía la necesidad de controlar a los distintos ámbitos de gobierno de diferentes adscripciones partidistas, misma que fue compartida y apoyada por la Cámara de Diputados. Sedesol la considera un esfuerzo por hacer coincidir estas gestiones con los conocimientos técnicos disponibles en el mundo académico; pero reconoce que se hizo con demasiado apresuramiento, se estructuró poco, y faltó mayor estudio y planeación.

Entre los legisladores y asesores del Congreso entrevistados las opiniones son variables: desde quien reconoce no saber específicamente por qué se tomó la decisión de evaluar, quien considera que la iniciativa correspondió a la CPCP de la LVII Legislatura para mejorar la asignación del gasto con base en los informes, quien opina que se trató de tener una mayor fiscalización de los recursos públicos federales para evitar que se utilizaran con fines electorales, y quien considera que su objetivo era contar con un diagnóstico de desempeño que diera continuidad al sistema de indicadores de evaluación iniciado durante el gobierno del licenciado Zedillo o con la ampliación de funciones de la ASF para la realización directa de evaluaciones.

Algunos académicos de relevancia consideran que es una razón técnica, la crisis fiscal, la que ha venido preocupando desde los ochenta a la SHCP y que,

en su afán por controlar el gasto gubernamental, ha querido realizar evaluaciones que le permitan recortar el gasto social improductivo, justificando a los ojos de la sociedad la reducción de programas y servicios públicos.

La variedad de razones expuesta induce a pensar que tanto funcionarios de las secretarías de Estado como legisladores, e incluso entre los últimos dependiendo de los partidos de adscripción, tuvieron motivos diferentes para aprobar esta medida, que sigue suscitando ciertos temores entre los responsables de los programas evaluados. Se trataría de una conjunción de voluntades en torno a una decisión pública, como medio, que sirve al logro de fines diferentes de los distintos grupos de actores, posibilidad perfectamente explicable bajo el modelo de políticas públicas de Lindblom (1959).

Por otra parte, no han quedado claros los criterios utilizados para sujetar o no un programa a reglas de operación y, en consecuencia exigir su evaluación. Esto también aparece incluido en el proyecto de presupuesto enviado por el Ejecutivo al Congreso, y los legisladores entrevistados no creen que en este momento la Cámara de Diputados pudiera intentar su modificación. Dicho proyecto se analiza en el marco de las comisiones y es aprobado por el pleno de Cámara; pero, en la medida en que ésta no lo analice a fondo constituye un medio para evitar la evaluación de un programa no muy visible que haya recibido cuestionamientos previos importantes. El mecanismo ha dejado fuera desde el inicio a diversos programas a cargo de institutos o fideicomisos y tampoco evalúa los programas a cargo de instituciones privadas que reciben recursos públicos.

En todo caso, actualmente los legisladores coinciden en que la evaluación no está sirviendo para lograr una mayor transparencia del ejercicio de recursos públicos y para mejorar la asignación del gasto, el cual sigue aprobándose bajo presión política y de tiempo, sin que los diputados revisen los informes y resúmenes ejecutivos de las evaluaciones realizadas, actividad que incluso ha rebasado la capacidad de la CACP.

## Organización y gestión del sistema

Formalmente, el sistema de evaluaciones externas se organiza y gestiona con base en la normatividad revisada en el capítulo anterior. Sin embargo, más allá de dicha organización y funcionamiento formal, nos interesa, retomando las ideas de Herbert Simon (1972) y diversos autores franceses que rescatan la informalidad en los estudios de regulación social (véase capítulo 1), analizar, la manera en que se relacionan las diversas instancias intervinientes en el proceso, la forma en que se aplican efectivamente los procedimientos

formales, las posibles negociaciones implícitas, los intereses que mueven a los actores participantes y sus consecuencias.

### *Organización y gestión formal*

En el diagrama 2 se relacionan las principales instancias que, de acuerdo con el marco jurídico, participan en el proceso de evaluación de programas sociales:

- La Cámara de Diputados, a propuesta del Ejecutivo, da origen al proceso de evaluación a partir del año 2001.
- Lo anterior afecta a todas las secretarías que desarrollan programas con recursos fiscales sujetos a reglas de operación, en particular, a la Sedesol, objeto de nuestro estudio.
- En 2002, la Sedesol contó con 31 programas, 13 de ellos operados en forma directa por su Dirección General de Políticas Sociales y otros 18 sectorizados en empresas y organismos desconcentrados.
- Cada uno de los anteriores cuenta con una oficina encargada de efectuar las licitaciones, contratos, control de la entrega de informes convenidos, relacionarse con las instituciones y los equipos de investigadores que realizan las evaluaciones, enviar los informes finales a la Sedesol y autorizar el pago al evaluador.
- Por último, la Sedesol realiza comentarios a los evaluadores y entrega los informes a la Cámara de Diputados, SHCP y SFP.

El sistema integrado por todas estas instancias constituye el contexto inmediato del trabajo que desarrollan los evaluadores externos.

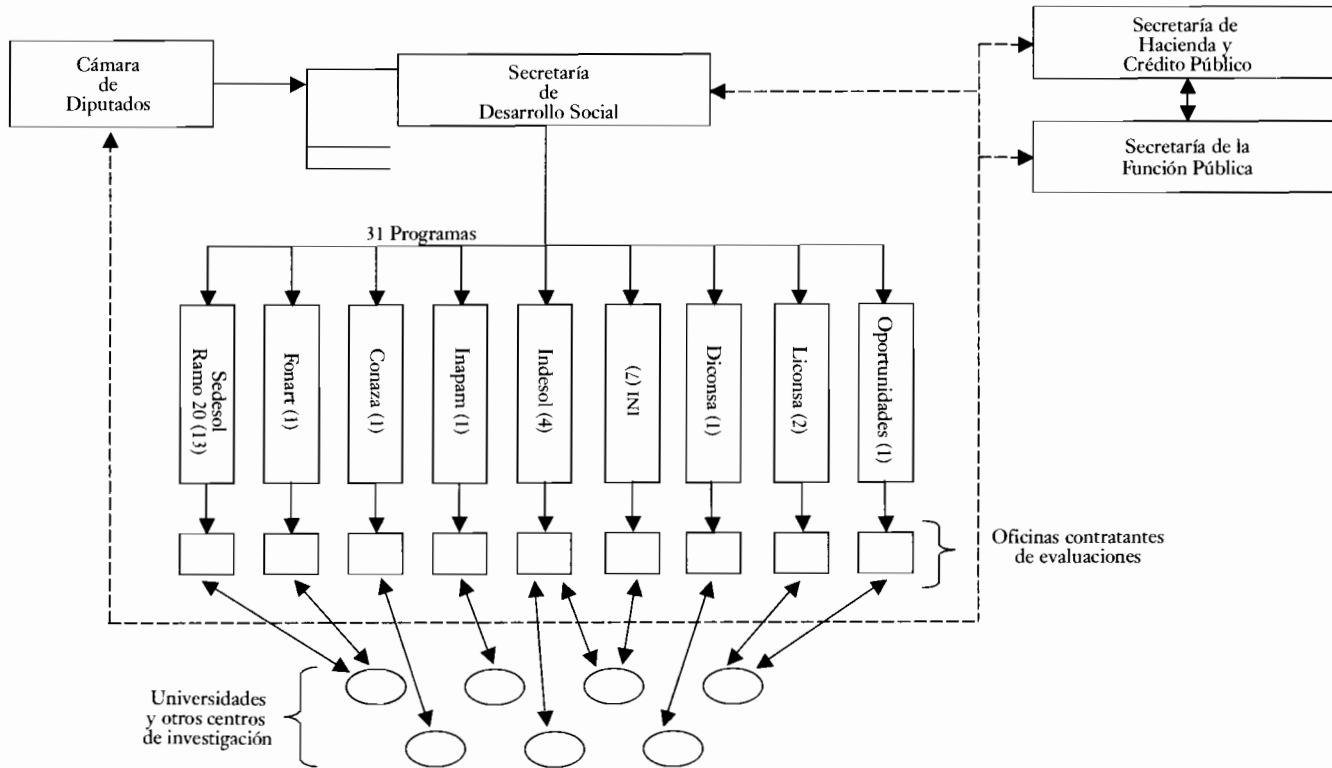
### *Funcionamiento real del sistema*

En la práctica, el Congreso no realiza la función prevista en la decisión formal, y la SHCP y la SFP cumplen un papel muy restringido, por lo cual el proceso queda a cargo de la Sedesol, quien es la principal destinataria real de sus resultados.

Además de conocer el *grado de participación de cada instancia*, se considera importante identificar la cantidad de interacciones que se realizan entre las *múltiples instancias que se relacionan* para llevar a cabo los procesos de evaluación y la calidad de las mismas en opinión de los actores participantes en el proceso. Así surgen como las instancias con mayores vínculos: la Sedesol

DIAGRAMA 2

# INSTANCIAS RELACIONADAS CON LA EVALUACIÓN DE PROGRAMAS SOCIALES



(Dirección General de Evaluación), los funcionarios administrativos y de áreas sustantivas de las instancias operadoras de programas, y los evaluadores externos.

Entre sus relaciones aparece como mejor calificada la que mantienen las áreas sustantivas de los programas con los demás actores; el resto recibe calificaciones medias, resaltando que las mayores dificultades surgen entre los evaluadores externos y los funcionarios de Sedesol y los administrativos del programa (véase gráfica 2). Al describir la última relación un evaluador manifestó textualmente: “La Sedesol actúa como si los evaluadores fueran culpables hasta que no demuestren su inocencia”, aclarando que como equipo evaluador su relación era buena, pero desde el punto administrativo el área contratante (Sedesol) presentaba muchas deficiencias y retrasos que postergaban la liberación del pago final.

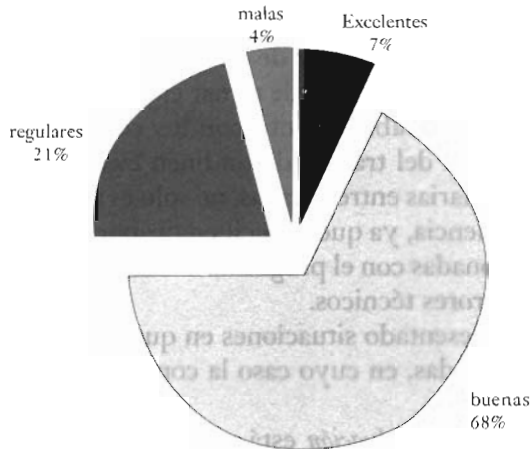
En casos en que la relación con Sedesol fue calificada como “regular”, esto se debió fundamentalmente a la poca frecuencia en la interacción con algún evaluador, o bien a desacuerdos con las instancias operadoras en torno a las estrategias de evaluación; pero se reconoció que la situación estaba mejorando.

Se trata de resultados que confirman lo que podía anticiparse ya que las áreas sustantivas, responsables de la gestión de los programas evaluados, son las que más cuidado ponen en sus relaciones con el resto; mientras que los evaluadores e instancias administrativas son los que mayores conflictos de intereses pueden enfrentar en torno a recursos, tiempos, requisitos administrativos, etcétera.

Los *términos de referencia*, que establecen las características del producto esperado de la evaluación, así como los requisitos de participación, responden a las reglas de operación del programa y al Acuerdo. Su redacción es casi idéntica en todos los casos, presentándose muy poca consideración de los aspectos específicos de cada programa, y dejando algunas ambigüedades o induciendo confusiones, al mismo tiempo que se muestra muy ambiciosa en casos que no lo justifican. La respuesta de los evaluadores a esta situación ha sido ofrecer lo que se pide, a pesar de que no podrán satisfacerlo, e intentar llenar el punto con un discurso bastante vacío de contenido. Es un juego en que se piden imposibles, se responde como si se hubieran cubierto y quienes hacen la revisión no exigen congruencia. En este sentido, un evaluador externo afirmó que se trata de “...un mercado de compradores que no saben lo que piden y de proveedores que no saben lo que pueden”.

Los *procedimientos de convocatoria y selección* de evaluadores se han llevado a cabo con relativa normalidad. De acuerdo con la Ley de Adquisiciones se

GRÁFICA 2  
CLASIFICACIÓN DE RELACIONES ENTRE ACTORES



debe realizar un concurso público para contratar a organizaciones privadas, salvo los casos de excepción revisados en el capítulo anterior. En la práctica de la evaluación se ha venido aplicando el procedimiento de invitación a cuando menos tres organizaciones, salvo que la evaluación sea efectuada por una organización del sector gubernamental y se considere que no existe otra con infraestructura, personal y capacidad suficiente para realizarla, en cuyo caso se procede a la adjudicación directa,<sup>67</sup> con base en el artículo 1o. de la Ley de Adquisiciones que, como se mencionó en el capítulo anterior, exime de la aplicación de la misma cuando los contratantes son dependencias o entidades de la administración pública.

La adjudicación directa se ha aplicado en el caso de las evaluaciones realizadas a Progresía/Oportunidades desde el año 1997 al 2003. También se habría intentado aplicarla nuevamente para la adjudicación directa de las evaluaciones multianuales programadas, provocando un profundo descontento en evaluadores externos que desean participar en su concurso. Ellos piensan que estos procedimientos simplificados se usan para evitarse más trabajo administrativo y que se corre el riesgo de que además existan negociaciones implícitas del tipo “tú me pagas bien, yo te evalúo bien”.

<sup>67</sup> Según información recogida en entrevistas a funcionarios de Sedesol y evaluador externo de Oportunidades.

En el caso del procedimiento por invitación se forma una comisión, integrada por miembros de las áreas jurídica y de evaluación de la Sedesol y de la instancia responsable del programa, que analiza las propuestas recibidas que cubren todos los aspectos solicitados y elige la más económica, de acuerdo con el criterio especificado en la Ley de Adquisiciones. Esto provoca serios problemas ya que la comisión no puede tomar en cuenta criterios vinculados a la calidad estimada del trabajo, junto con los costos presupuestados, ni garantizar la continuidad del trabajo de un buen evaluador. Como lo manifiesta una de las funcionarias entrevistadas, no sólo es un problema de calidad sino también de pertinencia, ya que se reciben propuestas de equipos formados en áreas no relacionadas con el programa a evaluar. En algunos casos esto también se presta a errores técnicos.

También se han presentado situaciones en que se reciben propuestas de organizaciones no invitadas, en cuyo caso la comisión *ad hoc* discute si se le permite o no participar.

Como el *mercado de la evaluación* está aún en proceso de conformación, también se presentan casos en que no hay instituciones interesadas en realizar la evaluación. En estas situaciones, después de dos convocatorias públicas o por invitación, se puede efectuar una adjudicación directa de la evaluación a alguna organización que se considere podrá cumplir las expectativas. Aún en estas condiciones, la contratación ha resultado difícil y, a veces, imposible.

Las grandes universidades y los investigadores de reconocido prestigio se resisten a participar por considerar que los recursos adjudicados son muy limitados y que se requiere realizar un excesivo papeleo burocrático. Por otro lado, existen muchos pequeños centros en crecimiento, interesados en obtener proyectos que les permitan continuar en el mercado y dispuestos a sacrificar parte de sus posibles ganancias con tal de obtener el contrato. Esto ha provocado que, en 2002 y 2003, los recursos disponibles para remunerar las evaluaciones realizadas sean bastante superiores (hasta 40 por ciento) a los efectivamente asignados.

En síntesis, la oferta de evaluadores es considerada limitada en cantidad y en conocimientos, en parte como consecuencia del desinterés de Conacyt por promoverlos y de la ausencia de programas de formación universitaria y diplomados en la materia. Por ello se ha propuesto certificar a quien tiene los conocimientos necesarios en evaluación para que las mismas aporten valor agregado al programa y sustenten adecuadamente sus juicios valorativos.

Como resultado de los procesos revisados, es posible afirmar que las asignaciones realizadas cubren, en términos generales, con los requerimientos establecidos en el Acuerdo y Términos de Referencia. Sin embargo, preocu-



pa que las excepciones a esta regla se produzcan en casos en que los presupuestos destinados a la evaluación resultan ser los más cuantiosos (por ejemplo Oportunidades). Lo anterior, unido a que el pago de la evaluación siempre es realizado por la misma instancia que opera el programa, deja dudas sobre la independencia de criterio del evaluador.

En términos económicos, la situación descrita caracteriza al mercado de la evaluación como un monopsonio (un solo comprador y múltiples oferentes) para cada programa, en que el servicio en cada caso debe ser diferenciado, lo que implica importantes costos de entrada al mercado para los oferentes potenciales. Así, no es posible evaluar de igual manera a Oportunidades, Abasto Social de Leche o Bienestar Social dadas sus profundas diferencias en cuanto a recursos, población atendida, objetivos, tipos de apoyo, etcétera.

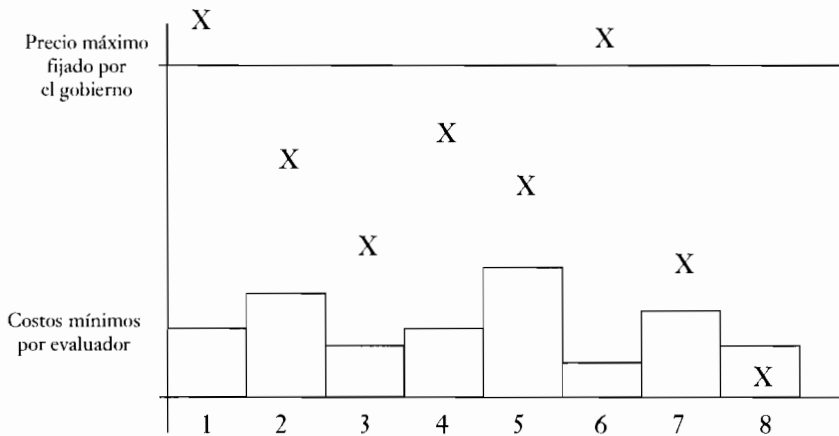
Obviamente, quien tiene un poder muy importante en este tipo de mercado es el comprador único (gobierno), lo que en el caso en estudio se traduce en su libertad para determinar unilateralmente el precio máximo que está dispuesto a pagar por cada evaluación, mismo que los oferentes saben que existe, pero no conocen oficialmente su monto (aunque regularmente el mismo se “filtra”). Esto acompañado de sus atribuciones exclusivas para seleccionar al evaluador con información casi “perfecta”, valorar su trabajo, autorizar su pago, etcétera, genera condiciones de profunda desigualdad entre las dos partes de ese mercado.

De esta forma, los oferentes se mueven entre dos límites para hacer sus propuestas económicas: el límite superior ya mencionado y un límite inferior determinado por su estimación de costos para realizar el trabajo (incluyendo retribuciones mínimas aceptables para el equipo de trabajo). Además la franja intermedia resultante se caracteriza por una alta incertidumbre: por un lado, el límite superior no es de conocimiento público y, aún si se llega a conocer, el oferente tiene que considerar la posibilidad de que otros competidores participen en el juego, en cuyo caso el contrato deberá otorgarse a quien, cumpliendo con los requerimientos de los Términos de Referencia, ofrezcan el menor precio. De esta forma, el oferente tiene que decidir entre múltiples combinaciones posibles de ganancia y riesgo, agravadas por el hecho de que si se equivoca no venderá sus servicios y tendrá que esperar un año para corregir su precio.

En consecuencia, un oferente puede verse enfrentado a cuatro situaciones distintas, en que una sola resulta ser una estrategia ganadora (véase gráfica 3):

- Hacer una propuesta económica muy elevada que lo deje fuera del juego (caso de los oferentes núms. 1 y 6 en la gráfica).

GRÁFICA 3  
MERCADO DE UNA EVALUACIÓN (MONOPSONIO)



Fuente: Elaboración propia.

- Realizar una propuesta económica aceptable pero enfrentarse a un competidor que ofrece menor precio y, en consecuencia, perder el concurso (casos núms. 2, 3, 4 y 5 en la gráfica).
- Realizar la propuesta económica aceptable y de menor precio y, en consecuencia, ganar el concurso (caso núm. 7 entre quienes cubren sus costos mínimos en la gráfica).
- Hacer una propuesta económica tan baja (por mala estimación de costos) que, en caso de ganar el concurso, no le permitiría cubrir sus costos mínimos.

En la práctica, lo anterior lleva a que algunas instituciones sacrifiquen el pago de honorarios de manera de tener mayores posibilidades de ganar el concurso. Los resultados han sido importantes recursos ahorrados por las instancias contratantes, pero demasiado escasos para lograr evaluaciones de alta calidad en la mayoría de los casos.

Puesto que esta última es un elemento que incide poco en el otorgamiento del contrato ya que sólo se verifica el cumplimiento de los requisitos mínimos, en la medida en que el número de oferentes aumente (actualmente son pocos todavía), el poder del comprador único crecerá, al mismo tiempo que los oferentes potenciales verán incrementarse los costos que implica intentar ingresar al mercado, ya que contarán con menos posibilidades de ganar los

concursos y, para aumentar las mismas, tendrán que reducir aún más sus aspiraciones en términos de retribución a su trabajo, lo que en el mediano plazo puede volver al mercado muy poco atractivo.

Otro mecanismo que permite ganar competitividad es la acumulación de contratos realizando economías de escala. Como puede verse en la gráfica, el límite inferior de la franja, correspondiente a los costos totales de producción de cada evaluador, es una línea quebrada que depende tanto de las mencionadas aspiraciones en términos de retribución como de la eficiencia interna del equipo de trabajo. Esta última puede incrementarse cuando los costos totales se aumentan, disminuyendo los costos medios, o sea distribuyéndolos entre un número aún mayor de acciones.<sup>68</sup>

Finalmente, habría que agregar los casos en que el demandante único decide que hay sólo un oferente capaz de brindar el servicio de evaluación, que además forma parte de la administración pública, por lo que le adjudica el contrato en forma directa, obviando el concurso público abierto o por invitación, en cuyo caso también el monoposonio se diluye, con grandes ventajas para el proveedor identificado y ninguna posibilidad de ingresar al mercado para el resto de los aspirantes.

En cuanto a la coordinación de las actividades de evaluación externa, en 2001, la Sedesol se encargó directamente de realizarla y seleccionó a las instancias académicas. Sin embargo, por haber sido el primer año en que se hacía este tipo de evaluaciones, algunos funcionarios entrevistados manifestaron que obtuvieron evaluaciones de mala calidad en términos de su diseño y aportaciones reales hacia los programas.

Debido a la experiencia de esta evaluación, la Cámara de Diputados solicitó se emitieran reglas claras para elevar su calidad, lo que originó la publicación del Acuerdo sobre los requisitos mínimos que deberían contener las evaluaciones que se realizaran por instancias académicas. También se decidió solicitar al Conacyt información de las instituciones que, por su nivel académico, cuerpo de investigadores, instalaciones y experiencia, fueran susceptibles de ser contratadas para realizar evaluaciones.

En ese año la cobertura de las evaluaciones no fue total: por lo menos en los programas a cargo de Liconsa y algunos del INI (Promoción de Culturas Indígenas, por ejemplo) la evaluación cubrió sólo algunos meses del año.

<sup>68</sup> Esto puede ejemplificarse en el caso de cuatro evaluaciones del INI que fueron realizadas por un mismo equipo: puesto que las poblaciones objetivo de estos programas eran iguales o muy similares, los evaluadores pudieron reducir costos al recoger información de gabinete, sistematizarla y analizarla, que resultó útil para más de un informe, así como aplicar más de una encuesta en una misma localidad (al eliminar gastos de transporte hacia otra localidad).

CUADRO 6  
PROGRAMAS SOCIALES EVALUADOS EN 2002

<i>Programas</i>	<i>Dependencia responsable</i>
<i>a. Sectorizados</i>	
1. Desarrollo Humano Oportunidades	Coordinación Nacional de Oportunidades
2. Abasto Social de Leche	Liconsa
3. Programa Tortilla	Liconsa
4. Abasto Rural	Diconsa
5. Promoción de Culturas Indígenas	INI
6. Desarrollo Económico Productivo de los Pueblos Indígenas	INI
7. Planeación	INI
8. Bienestar Social	INI
9. Reparación y Equipamiento de Albergues	INI
10. Operación de Albergues	INI
11. Promoción y Procuración de Justicia	INI
12. Programa de Vinculación y Transferencia Tecnológica, Plan Acuario y Programa para el Desarrollo de los Habitantes del Semidesierto	Conaza
13. Compras, Créditos, Concursos y Asistencia Técnica	Fonart
14. Expertos en Acción	Inapam
15. Coinversión Social	Indesol
16. Investigación para el Desarrollo Local	Indesol
17. Capacitación para el Fortalecimiento de la Sociedad Civil	Indesol
18. Fortalecimiento Institucional para la Gestión Social Municipal	Indesol
<i>b. Ramo 20 (Desarrollo Social)</i>	
19. Ahorro y Subsidios para la Vivienda Progresiva	Sedesol
20. Identidad Jurídica	Sedesol
21. Zonas de Alta Marginación e Indígenas	Sedesol
22. Microrregiones	Sedesol
23. Atención a Jornaleros Agrícolas	Sedesol
24. Desarrollo de Pueblos y Comunidades Indígenas	Sedesol
25. Empleo Temporal	Sedesol
26. Estatales por Demanda	Sedesol
27. Iniciativa Ciudadana 3 × 1	Sedesol
28. Jóvenes por México	Sedesol
29. Mujeres Jefas de Familia	Sedesol
30. Oportunidades Productivas	Sedesol
31. Superación de la Pobreza Urbana	Sedesol

Fuente: Elaboración propia a partir de los informes finales de evaluación, 2003, entregados por Sedesol.

En 2002 se evaluaron los 31 programas reportados (véase cuadro 6); el trabajo fue realizado por 21 *instituciones evaluadoras*. Como puede observarse, el caso de Oportunidades es excepcional, ya que confluyen para su realización diversas instituciones que manejan temas específicos y un importante conocimiento acumulado; por eso constituye la situación que la Dirección General de Evaluación quisiera alcanzar como meta en todos los programas.

Sólo dos instituciones participaron en la evaluación de cuatro programas.<sup>69</sup> En el primer caso se trata de programas dirigidos a población indígena, que coincide con el objeto de estudio del centro evaluador, lo que facilitó el trabajo.

En cambio en el segundo, se evaluaron diversos programas del Ramo 20, dirigidos a poblaciones distintas, con objetivos diferentes, etcétera, lo que aumenta el riesgo de no lograr buenas calidades en todas las evaluaciones.

La distribución de los evaluadores aparece concentrada en el centro del país con pocas excepciones. También es muy reducido el número de casos en que participan organizaciones de carácter privado.

El móvil que induce a la participación institucional revisada es fundamentalmente la obtención de experiencia y reconocimiento a su trabajo; sólo en casos excepcionales de montos muy elevados, éste mismo se transforma en un incentivo *per se* para la institución.<sup>70</sup>

Algo similar sucede con los coordinadores de equipos evaluadores, si bien en algunos casos reciben ingresos adicionales sustantivos. Sin embargo, hay que destacar la existencia de un grupo de asistentes de investigación y encuestadores para los cuales la evaluación llega a constituir una oportunidad única de obtener ingresos, en este contexto de insuficiencia y precarización creciente del empleo en el país. Se constató en varios casos que los últimos pueden llegar a impulsar a los coordinadores a participar.

Finalmente, hay que resaltar que los problemas de gestión administrativa aparecen como relevantes entre los cuatro funcionarios de instancias operativas entrevistados: cumplimiento de la normatividad, aprobación de trámites por parte de la contraloría interna, manejo presupuestal y administrativo de las instancias académicas, retrasos en las fechas de entrega, falta de experiencia en la redacción de términos de referencia, etcétera.

<sup>69</sup> En opinión de una funcionaria entrevistada, la concentración institucional aumenta cuando se analiza el número de evaluaciones realizado por cada organización para el conjunto de las secretarías de Estado.

<sup>70</sup> Estos varían entre el 10 y el 30 por ciento del monto contratado.

CUADRO 7  
PRESUPUESTOS DE LOS PROGRAMAS  
Y DE SUS EVALUACIONES, 2002

<i>Programas</i>	<i>Presupuesto ejercido por el programa (miles de pesos)</i>	<i>Costo de la evaluación* (miles de pesos)</i>	<i>Porcentaje Costo/presupuesto (porcentaje)</i>
1. Desarrollo Humano Oportunidades	5'965,616.70	28'881.89	0.48
2. Abasto Social de Leche	0.00	1'150.00	N.C.
3. Programa Tortilla	774,534.20	498.00	0.06
4. Abasto Rural	6'699,588.90**	2'080.00	0.03
5. Promoción de Culturas Indígenas	57,568.30**	244.95	0.50
6. Desarrollo Económico Productivo de los Pueblos Indígenas	367,899.10	665.00	0.19
7. Planeación	37,167.90	919.44	0.20
		<i>por los programas Nos. 7 a 10</i>	
8. Bienestar Social	17,042.50**		
9. Reparación y Equipamiento de Albergues	119,453.80		
10. Operación de Albergues	292,049.70		
11. Promoción y Procuración de Justicia	35,962.00	550.00	1.53
12. Programa de Vinculación y Transferencia Tecnológica, Plan Acuario y Programa para el Desarrollo de los Habitantes del Semidesierto	101,983.20	899.00	0.88
13. Compras, Créditos, Concursos y Asistencia Técnica	26,981.80	1'132.81	4.20
14. Expertos en Acción	50,219.60	1'000.00	1.99
15. Coinversión Social	116,170.00	150.00**	0.13
16. Investigación para el Desarrollo Local	15,497.60	221.00	1.52
17. Capacitación para el Fortalecimiento de la Sociedad Civil			
18. Fortalecimiento Institucional para la Gestión Social Municipal	24,665.90	700.00**	2.84
19. Ahorro y Subsidios para la Vivienda Progresiva	882,474.80	1'100.00	0.12
20. Identidad Jurídica	12,813.10	655.30	5.11
21. Zonas de Alta Marginación e Indígenas	101,578.70	568.40	0.56
22. Microrregiones	220,000.50	970.01	0.44
23. Atención a Jornaleros Agrícolas	172,978.90	848.58	0.49
24. Desarrollo de Pueblos y Comunidades Indígenas	824,682.80	780.33	0.09
25. Empleo Temporal	1'522,146.20	2,001.10 SCT 2,926.38 Sagarpa***	0.32
26. Estatales por Demanda	368,185.10	300.00	0.08
27. Iniciativa ciudadana 3×1	99,586.90	576.26	0.58
28. Jóvenes por México	140,052.70	800.00	0.57

<i>Programas</i>	<i>Presupuesto ejercido por el programa (miles de pesos)</i>	<i>Costo de la evaluación* (miles de pesos)</i>	<i>Porcentaje Costo/presupuesto (porcentaje)</i>
29. Mujeres Jefas de Familia	137,718.80	300.00	0.22
30. Oportunidades Productivas	910,079.20	725.00	0.08
31. Superación de la Pobreza Urbana	199,305.80	800.00	0.40
Totales	20'294,004.70	52,546.00	0.26
Costo promedio total		1,695.00	
Costo promedio excluyendo Oportunidades		788.80	
Costo promedio s/Oportunidades, Empleo Temporal y Abasto Rural		594.9	

\*Las instituciones evaluadoras retienen entre un 10 y un 30 por ciento por concepto de gastos de administración.

\*\* La información recibida de la fuente anterior presenta en algunos casos diferencias sustantivas respecto de la presentada en la Cuenta de la Hacienda Pública Federal 2002. ejemplo: presupuesto ejercido por Abasto Rural 568,000.00; Promoción de Culturas Indígenas 48,701.00; Bienestar Social 4,747.70 (en miles de pesos). Lo mismo sucede con el costo de las evaluaciones: coinversión Social 600.00; Capacitación y Fortalecimiento Institucional 700.00.-

\*\*\* Incluye el presupuesto del Centro de Información del PFT.

Fuente: Elaboración propia a partir de información aportada por el IFAI.

## Recursos para la realización de evaluaciones externas

En este apartado se consideran los recursos de toda índole: financieros, humanos, de información, etcétera.

### *Recursos financieros*

A continuación se analiza el monto de recursos financieros empleados en cada evaluación y también el presupuesto ejercido por el programa, para revisar si guardan relación y/o permiten ver el esfuerzo diferencial que realiza cada programa para ser evaluado (véanse cuadro 7 y gráfica 4). Los recursos disponibles para la evaluación los determina la dependencia sectorizada, o Sedesol en el caso de Ramo 20, y requieren de la autorización de la SHCP.

Como ya se mencionó, a veces hay diferencias importantes entre los recursos disponibles para la evaluación y los realmente asignados (hasta 40 por ciento menos), debido a que, de las propuestas recibidas que cumplen con los aspectos requeridos, se debe elegir la de menor costo, y ésta ha estado por debajo de los recursos autorizados. También se puede dar el caso de que alguna institución evaluadora haya estado investigando el tema y reciba recursos adicionales de otra fuente (Conacyt, Indesol, etcétera).

El caso de Oportunidades (evaluación 1 en la gráfica citada) resulta absolutamente extraordinario en sus montos (5,965.6 millones de pesos de presupuesto ejercido y 28.8 millones de pesos de costo de sus evaluaciones) y Abasto Rural (evaluación 3) los supera a todos en presupuesto (6,699.6 millones de pesos, según la información recibida de IFAT).

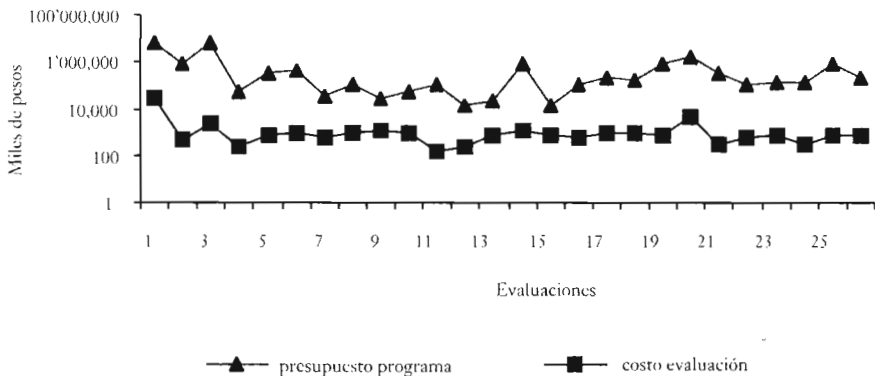
Los siguientes presupuestos más elevados corresponden a cuatro programas pertenecientes al Ramo 20: Empleo Temporal, Oportunidades Productivas, Ahorro y Subsidios para la Vivienda Progresiva y Desarrollo de los Pueblos y Comunidades Indígenas, con un promedio de 1,037 millones de pesos.

En cuanto al costo directo pagado por las evaluaciones, los más altos (después de Oportunidades) corresponden a dos programas de Ramo 20 y cuatro sectorizados: empleo.

Temporal, Abasto Rural y Abasto Social de Leche, programas de Fonart, Ahorro y Subsidios para la Vivienda Progresiva, y Expertos en Acción, con un promedio de 1.9 millones de pesos por evaluación.

En la relación entre costo de la evaluación y el presupuesto del programa evaluado (véanse cuadro 7 y gráfica 5) sobresalen por su elevado nivel Identidad Jurídica (5.11 por ciento) y los programas de

GRÁFICA 4  
PRESUPUESTO DE LOS PROGRAMAS Y COSTO DE LAS EVALUACIONES



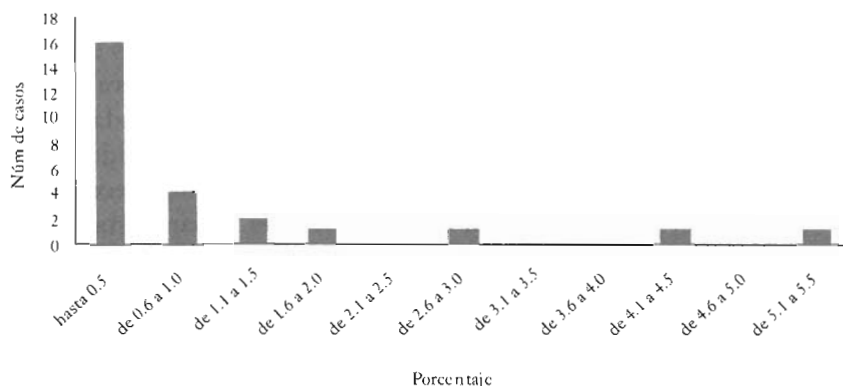
Notas: Gráfica realizada en escala logarítmica dada la gran dispersión de los datos.

Los datos de los programas núms. 7 a 10 y 17 y 18, que no se disponen en forma desglosada, se representan en conjunto.

Los datos del programa 2 (Abasto Social de Leche) no se representaron debido a que su presupuesto fue cero y el logaritmo correspondiente (-) carecería de sentido aplicado y quedaría fuera del cuadrante graficado.



GRÁFICA 5  
PORCENTAJE COSTO/PRESUPUESTO



Fonart (4.19 por ciento); mientras en el extremo opuesto se localizan 16 evaluaciones entre las que destacan: Abasto Rural (0.03 por ciento), Tortilla (0.05 por ciento) y Estatales por Demanda y Oportunidades Productivas (0.08 por ciento).<sup>71</sup>

Lo anterior muestra que no existe ninguna relación constante entre las dos variables analizadas, lo que ya se podía deducir de la ausencia de similitud en el trazado de sus curvas en la gráfica previa.

Estadísticamente, podemos descartar la existencia de algún tipo de correlación, incluso eliminando los tres programas cuyos datos se muestran más alejados del resto (Oportunidades, Abasto Rural y Empleo Temporal), ya que obtenemos un coeficiente de correlación<sup>72</sup> de 0.21 (lejano al valor de 1 o -1 que nos indica correlaciones directas o inversas altas) y una R2 de Pearson de 0.089 (que muestra que sólo una proporción muy pequeña de la varianza de la variable “costos de las evaluaciones” puede ser atribuida a la varianza de la variable “presupuesto de los programas”).

También eliminando los tres programas extremos, esto se puede observar al realizar una gráfica de datos, que muestra una gran dispersión de los mismos, en lugar de un agrupamiento en torno a una curva demostrativa de

<sup>71</sup>Mientras el promedio de esta relación alcanza el 0.26 por ciento en Sedesol, en el BM varía entre el 2 y 3 por ciento.

<sup>72</sup>El coeficiente de correlación se calcula como  $\rho(x,y) = \frac{\text{Cov}(x,y)}{\delta x \cdot \delta y}$

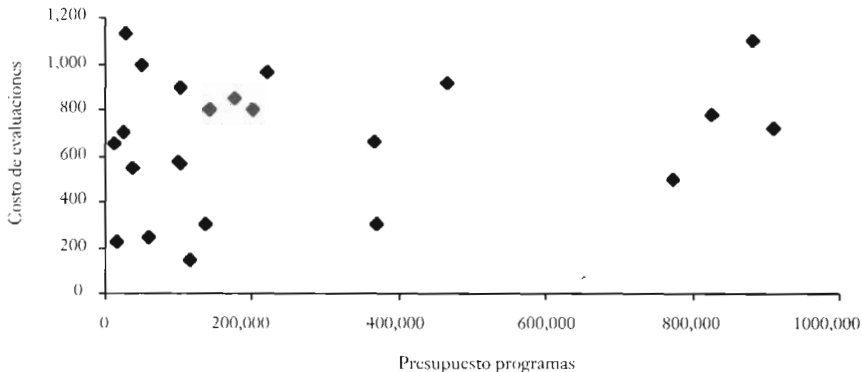
la existencia de algún nivel de correlación. En este caso la regresión mostraría más bien una recta casi horizontal (véase gráfica 6), donde el costo de la evaluación no depende del presupuesto e indica el promedio general de 594,900.00 de pesos.

En general, los evaluadores y funcionarios de los programas entrevistados consideran que los recursos financieros son insuficientes para el trabajo de campo que realizan, sobre todo en comunidades muy alejadas unas de otras, con excepción de Oportunidades, en que la mayoría considera excesivos los recursos disponibles. Algunos mencionan incrementos necesarios del 30 y 40 por ciento, pero consideran que esto debería variar bastante según los programas.

En cuanto a los legisladores y miembros de la ASF, por lo general desconocen los montos pagados para la realización de evaluaciones externas. Algunos se muestran sorprendidos al conocer sus montos y afirman que la idea de la CPCP era que éstas no rebasaran el 10 por ciento de los gastos administrativos del programa. Hay acuerdo entre ellos en el sentido de que no se establecieron criterios para definir el precio a pagar en cada evaluación.

Finalmente, también habría que considerar los costos indirectos de las evaluaciones, relativos al tiempo y esfuerzos invertidos por personal de la Sedesol y de las instancias ejecutoras de los programas para la realización de las evaluaciones. Dadas las dificultades para medir estos costos, debido a las múltiples actividades que desarrollan las 14 personas dedicadas a esta actividad en la Dirección General de Evaluación de Sedesol y la persona que se responsabiliza del trabajo en las instancias ejecutoras, así como sus colabo-

GRÁFICA 6  
DISPERSIÓN DE DATOS COSTO/PRESUPUESTO  
(En miles de pesos)



radores, el punto no será cubierto por la presente investigación, pero se puede estimar en aproximadamente 200,000.00 pesos en promedio por programa.

### *Recursos humanos*

En cuanto a la calificación y experiencia de los recursos humanos del sistema de evaluación, se comenzará por quienes se vinculan a ella a nivel central (Sedesol) y los funcionarios de las instancias operadoras que se encargan de la contratación de evaluadores, para continuar con las características de los equipos conformados por los últimos y finalizar con los destinados por otras secretarías relacionadas con la evaluación y el Congreso.

Entre los primeros, destaca el trabajo realizado en la Subsecretaría de Prospectiva, Planeación y Evaluación que, a partir de abril de 2002, fue ocupada por Miguel Székely Pardo, autor de diversos trabajos de evaluación y ex miembro del área social del BID. En opinión de Boltvinik (2002), se trata de un economista experto en análisis empíricos sobre pobreza, distribución del ingreso y sus relaciones con las condiciones generales de la economía, admirador del diseño y funcionamiento de Progresá.

La Dirección General de Evaluación, perteneciente a la Subsecretaría anterior, cuenta con unas 14 personas. En opinión del funcionario entrevistado, las mismas resultan insuficientes para profundizar en el análisis de cada programa. Manifiesta que se está privilegiando la evaluación de impacto económico, por lo que casi todos los miembros de la dirección de evaluación cuentan con formación en esa disciplina, con posgrados en instituciones extranjeras de excelencia, en áreas de econometría, evaluación de programas sociales, políticas públicas, etcétera. En particular, en 2002, todos los directores eran economistas, salvo el caso de un administrador público. Su formación explica el énfasis econométrico que parece imponerse como criterio de calidad. La mayoría de esos especialistas cuenta con experiencias relevantes de trabajo previo, aunque no necesariamente en evaluación. Lo anterior resulta consistente con la afirmación de un evaluador externo que manifiesta que a dicha dirección le ha faltado experiencia para organizar todo el sistema, aunque reconoce que ha avanzado mucho.

En la Dirección General Adjunta de Análisis Territorial, encargada de la estrategia transversal Microrregiones, se cuenta con unas ocho personas que apoyan su trabajo. La evaluación que se pretende realizar en esta dirección general presenta un enfoque sustantivamente distinto al anterior (enfoque territorial, impacto agregado de la inversión en infraestructura y ausencia de sesgo economicista) que, en parte, se relaciona con la formación diferente de quien la dirige: sociología y comunicaciones.

Concretamente, las cuatro personas entrevistadas a nivel central de la Sedesol (excluyendo a su director general) tenían los siguientes estudios:

<i>Licenciatura</i>	<i>Institución</i>	<i>Posgrados</i>	<i>Institución</i>
*Economía	nacional	Economía	nacional
*Administración Pública	nacional	Políticas Públicas	extranjera
*Sociología y Comunicación	extranjera	Desarrollo Regional	extranjera
*Psicología	nacional	Políticas Públicas	nacional

Ninguno de ellos cuenta con estudios formales especializados en evaluación, sino que su formación proviene parcialmente de sus estudios de posgrado, cursos, estancias y experiencias de trabajo, en algunos casos vinculadas a organismos internacionales. Su antigüedad en el cargo varía entre uno y cuatro años y en la Sedesol entre uno y 12. En un contexto en que, como ya se mencionó, predominan ampliamente los economistas, esta diversidad nos permitió contar con distintos puntos de vista.

Entre los responsables de la vinculación con los evaluadores en cada programa, contamos con la información y opiniones de cuatro de ellos:

<i>Licenciatura</i>	<i>Institución</i>	<i>Posgrados</i>	<i>Institución</i>
*Economía	ND	No	ND
*Sociología	extranjera	Economía	nacional
*Ciencias de la Comunicación	ND	Ciencias de la Comunicación	ND
*Actuaría	nacional	Estadística	extranjera

El primero cuenta con una experiencia previa en la evaluación de un proyecto; la segunda tiene amplia experiencia práctica en consultoría; la tercera tomó el curso sobre modelos econométricos impartido en Sedesol durante 2003; y la cuarta cuenta con especialidad en análisis de la información. No tienen estudios especializados en evaluación, sino como parte de sus posgrados o cursos intensivos. Sus antigüedades en el cargo varían entre uno y tres años, mientras en la institución han permanecido entre dos y cinco años.

A pesar de estas variadas formaciones y experiencias, en opinión de algunos evaluadores externos, no todos los responsables de la vinculación con evaluadores están suficientemente capacitados para calificar las propuestas técnicas y guiar la asignación de los contratos, que si bien se revisan en

la comisión *ad hoc*, esto se hace generalmente sobre la base de un resumen comparativo de ventajas y desventajas de las propuestas, preparado por los propios funcionarios del programa.

Similar opinión tiene la ASF que recomienda mejorar su capacitación para: producir buenos términos de referencia, simplificar las reglas de operación para facilitar a los evaluadores el acceso a los datos del programa, y mejorar la estructura de incentivos para contratar y supervisar las evaluaciones.

Una vez más el caso de Oportunidades sería la excepción: cuenta con un equipo interdisciplinario, compuesto por actuarios y matemáticos, especialistas en política social, economistas y sociólogos, que dominan el idioma inglés.

Para la Dirección General de Evaluación de la Sedesol los equipos de evaluadores externos deberían contar con experiencia en medición y evaluación, tanto con conocimientos de métodos cuantitativos como cualitativos e incluir entre sus miembros especialistas en muestreo, análisis estadístico y trabajo de campo. También estima importante que estos tengan conocimiento previo de los programas, ya que el tiempo disponible para la evaluación es muy limitado y familiarizarse con los programas insume una buena parte de él.

Entre los funcionarios de los programas evaluados encontramos algunos altamente satisfechos con la conformación interdisciplinaria y los resultados del equipo evaluador y otros que aún enfrentan dificultades como: ausencia de disciplinas necesarias en el equipo, perfiles dispersos en las diversas instancias participantes en la evaluación, falta de experiencia, ética, profesionalismo y compromiso para llevar a cabo buenas evaluaciones.

Con base en la información que ofrecen los informes de evaluación y la recogida mediante entrevistas, la composición de los equipos en 2002 resulta muy variada. Si bien la información disponible es insuficiente para realizar juicios categóricos; sin embargo, la relativa a 18 programas muestra:

- Amplia formación interdisciplinaria en ocho casos: Oportunidades, Abasto Social de Leche, Tortilla, Promoción de Culturas Indígenas, Planeación, Bienestar Social, Reparación y Equipamiento de Albergues y Operación de Albergues.
- Inadecuada formación disciplinaria para el programa evaluado en dos casos: Jóvenes por México y programas de Fonart.
- Muy buen nivel de formación del equipo en Abasto Social de Leche, Tortilla y Oportunidades, seguidos de las evaluaciones de programas a cargo del INI y Fonart.
- Algún tipo de experiencia reconocida en trece casos.

Si la información estuviera completa se tendría que concluir que sólo siete programas (Oportunidades, Abasto Social de Leche, Tortilla, Planeación, Bienestar Social, Reparación y Equipamiento de Albergues y Operación de Albergues) cuentan con equipos sólidos. Considerando que no resulta tan obvio que un trabajo de evaluación logre mejor calidad por estar hecho por profesionales con nivel de doctorado ni que esto deba regir para todos los tamaños de programas, se podría agregar un pequeño programa más (Promoción de las Culturas Indígenas).

Finalmente, en las entrevistas realizadas a los miembros de equipos se identificaron los siguientes perfiles:

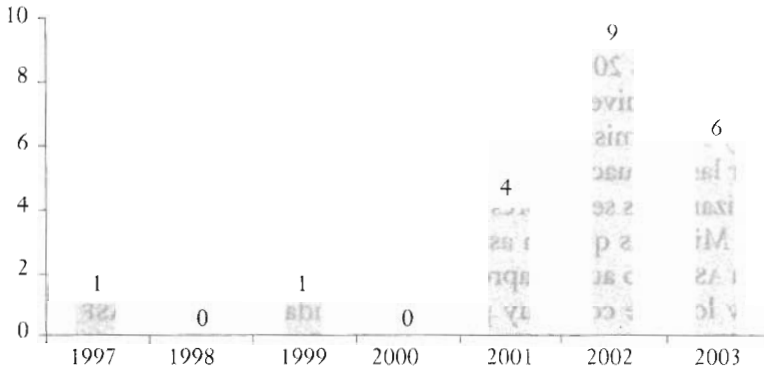
<i>Licenciatura</i>	<i>Institución</i>	<i>Posgrados</i>	<i>Institución</i>
*Economía	nacional	Maestría Administración Pública	nacional
		Doctorado Ciencias Económicas	nacional
*Sociología	nacional	Doctorado Sociología	nacional
*Nutrición	ND	Maestría Economía Salud	nacional
*Administración Pública y Ciencias Políticas	nacional	Maestría Planeación y Desarrollo	extranjera
		Estudios Doctorado Desarrollo Económico	extranjera
*Trabajo Social	nacional	Maestría Antropología	nacional
*Antropología	nacional	Doctorado Antropología	nacional

Se observa que tampoco los evaluadores cuentan con una formación específica en evaluación; sin embargo, dos de ellos incluyeron cursos sobre el tema dentro de sus programas de maestría y doctorado. Destaca por su formación el evaluador que realizó estudios en el extranjero.

Los anteriores han trabajado en la institución que representan entre tres y quince años (con un promedio de 11 años) y en su puesto actual entre dos y quince años (con una media de 6.6 años); lo que muestra que la amplia mayoría cuenta con gran experiencia.

En el área de evaluación, dicha experiencia se resume en la gráfica 7, que evidencia una débil experiencia previa al año 2001 (sólo dos evaluaciones), año en que inicia el proceso de evaluaciones generalizadas y en que dos de los cinco evaluadores entrevistados realiza el trabajo en cuatro programas. En 2002, el número de evaluaciones nacionales realizadas aumentó a nueve y participaron cuatro de los cinco evaluadores; en 2003, la situación se man-

GRÁFICA 7  
NÚMERO DE EVALUACIONES REALIZADAS  
POR LOS ENTREVISTADOS



tiene aproximadamente igual.<sup>73</sup> Esta tendencia se muestra consistente con la evolución en el número total de evaluaciones realizadas por año y muestra una experiencia promedio de 4.4 evaluaciones por institución en los tres últimos años, o sea 1.5 por evaluador y por año.

En opinión del director de metodología de la evaluación de Sedesol, la conformación de equipos sólidos en cuanto a su formación y experiencia coincide con las evaluaciones de mejor calidad: Oportunidades, Abasto Social de Leche, Tortilla y Abasto Rural.

Por otra parte, *los recursos humanos destinados a la revisión de los informes en la SHCP y SFP* son poco numerosos (unos cinco o seis) y sólo dedican parte de su tiempo a esta actividad.

Es importante resaltar que los miembros de la Subsecretaría de Prospectiva, Planeación y Evaluación manifestaron el interés de ésta en fortalecer la formación de recursos humanos en materia de evaluación con un perfil de economistas con doctorados en estadística, econometría, políticas públicas, etcétera, y mediante capacitación con enfoque econométrico.<sup>74</sup>

La Dirección General de Evaluación también organiza seminarios periódicos. El director de metodología de la evaluación consideró que estos han

<sup>73</sup> CIFSAS evaluó un solo programa debido a la reestructuración del INI en Comisión Nacional para el Desarrollo de los Pueblos Indígenas.

<sup>74</sup> Es el caso del curso de actualización dirigido principalmente a evaluadores institucionalmente invitados, desarrollado en 2003 por la propia Sedesol, con participación de profesores de FUA, en que se reconoce que hubo evaluadores interesados, pero no fueron numerosos.

ayudado a mostrar nuevas técnicas y formas de análisis, distintas a las que aplica el evaluador, para motivarlo a investigar más sobre el tema, y que también ha sido de beneficio interno a la Sedesol, ya que asiste personal de otras direcciones. Reconoció que su impacto no ha sido evaluado; la observación personal realizada en algunos eventos muestra que la afluencia del público es muy limitada (unas 20 personas por sesión).

Finalmente, a nivel del Poder Legislativo, se afirma que la Cámara de Diputados y sus comisiones no disponen de personal suficiente ni capacitado para revisar las evaluaciones, salvo en el caso de la ASF, por lo que se propone profesionalizar a sus servidores públicos y poner en práctica un servicio civil de carrera. Mientras que un asesor de un partido de oposición (PRD) manifestó que la ASF sólo audita aproximadamente el 5 por ciento de los recursos ejercidos, y lo hace con muy poca oportunidad; la propia ASF señaló que fiscaliza el 2 por ciento de las acciones realizadas por el Ejecutivo, lo que considera un buen nivel de ese indicador, ya que EUA verifica el 5 por ciento, Canadá el 6 por ciento, Reino Unido el 6.5 por ciento y Finlandia, que carece de corrupción, el 10 por ciento.

Un último comentario lo merece un rasgo importante de toda persona vinculada a este proceso, sobre el que no se hace consideración alguna: el moral. Como planteamos en el capítulo 2, la credibilidad de una evaluación para la opinión pública depende probablemente más de rasgos de honestidad y honradez probada de los evaluadores que de las características técnicas de su trabajo. Guardadas todas las proporciones (ya que en este caso se emiten recomendaciones y no sentencias), su trabajo presenta características similares a las de un miembro del Poder Judicial o un organismo como el Instituto Federal Electoral, en la medida en que le toca juzgar una gestión, motivo por el que cuidar sus condiciones éticas resulta relevante. En el momento actual existen dos mecanismos legales que no garantizan plenamente la independencia de criterio de los evaluadores externos: uno, que el costo de las evaluaciones sea autorizado y pagado por el propio programa evaluado; y otro, que exista la posibilidad de adjudicar presupuestos cuantiosos por la vía de la asignación directa.

Este tema de la ética del evaluador tendrá que ser especialmente cuidado, conjuntamente con la calidad, al momento de que se integre el nuevo CNE, organismo creado en sustitución del instituto autónomo que se había propuesto durante el debate de la LGDS. Varios legisladores entrevistados fueron partidarios de la creación de una instancia que dependa del Congreso o de la ASF para sustituir a este Consejo en la coordinación de las evaluaciones.



### *Otros recursos*

Otro recurso escaso muy importante para la realización de este trabajo lo constituye el tiempo disponible y la calendarización de las actividades e informes. En 2002 se dispuso de aproximadamente siete meses para la realización de las evaluaciones, tiempo que se considera insuficiente por los evaluadores y muy restringido en comparación con el que suelen dedicar los países que presentan un mayor grado de desarrollo de estas experiencias, revisados en el capítulo 4. Adicionalmente, suelen presentarse retrasos sobre la marcha, como la provocada por las dificultades generadas ese año en el diseño de la muestra. En 2003, algunos evaluadores informaron que el tiempo se redujo a cuatro meses debido a que se suscitaron nuevos problemas.

También se exige en algunos programas la entrega de un informe de resultados parciales en el mes de diciembre, que luego se tiene que modificar totalmente al completarse el trabajo de campo y redactarse el informe final del siguiente mes de marzo, duplicando esfuerzos.

Como ya se mencionó, en opinión de uno de los responsables de la contratación de evaluadores externos, las dificultades aumentarán en el futuro, debido a las modificaciones en el calendario que al respecto prevé la nueva LGDS.

Finalmente, no se perciben problemas en cuanto a equipo técnico y paquetería disponible por los evaluadores; en cambio, otro recurso muy escaso lo constituye la información estadística y documental disponible en la operación de los programas, que se revisará en el capítulo siguiente.

### Comentarios finales

El presente capítulo ha aportado la revisión de las condiciones contextuales, favorables o desfavorables, en que el proceso de evaluación de programas sociales se lleva a cabo, tratando de comprender las razones que indujeron a su implantación, cómo se planea complementarlo, la forma en que realmente funciona, y con qué recursos cuenta.

En términos generales, encontramos un panorama de alta coincidencia política en la decisión de echar a andar el proceso, acompañada por algunas resistencias culturales, en que los recursos institucionales (presupuestales y de información) presentan en la mayoría de los casos limitaciones importantes para el desarrollo de trabajos de calidad y un mercado de evaluaciones aún en formación, contando con una oferta, también restringida, de capacidades especializadas en el tema. Se observa también un funcionamiento del sistema relativamente apegado a su concepción formal.

Sin embargo, es necesario subrayar la identificación de múltiples problemas que afectan los trabajos de los evaluadores en materia de:

- Factibilidad de la evaluación: afectada por una cultura de la evaluación pobremente desarrollada, un Poder Legislativo poco presente como contrapeso de las acciones del Ejecutivo y una ausencia de criterios para definir los programas sujetos a reglas de operación, y por lo tanto, los obligados a realizar evaluaciones externas.
- Costos de la evaluación: sin criterios para definir los techos financieros disponibles por cada programa; indirectamente incrementados por la existencia de una oficina encargada de la relación con los evaluadores en todos ellos.
- Calidad de la evaluación: diseño de programas con objetivos poco claros y medibles, ausencia de Términos de Referencia *ad hoc* a cada programa, recursos financieros no relacionados con la importancia del programa, recursos humanos con insuficiente calificación en el tema, acumulación de contratos por algunos equipos de evaluadores, riesgo de pérdida de independencia en el criterio de valoración empleado, alto costo de entrada de evaluadores al mercado, evaluaciones adjudicadas sin licitación pública previa, y dificultades administrativas en la relación de los evaluadores y Sedesol.

Una vez presentada la información del contexto en que el proceso en estudio se desarrolla, en el siguiente capítulo se procederá a realizar la metaevaluación transversal de las 31 evaluaciones aplicadas a dichos programas durante el año 2002, y un breve análisis longitudinal de algunas de ellas.

## Capítulo 11

# Métodos, calidad y costos de las evaluaciones externas

### Introducción

A DIFERENCIA DE LO planteado por Duran (2002) para el contexto francés en el sentido de que el análisis del contexto es más relevante que el del método de evaluación, la presente tesis considera aún muy importante la inclusión de este aspecto para el nivel de desarrollo de nuestro país en la materia. Se trata de investigar si las actividades desarrolladas en el proceso de investigación son congruentes entre sí, con la metodología adoptada y con los objetivos perseguidos por la evaluación, y suficientes para fundamentar la valoración de la gestión del programa.

Por lo tanto, el presente constituye uno de los capítulos centrales de este trabajo. En él se procede a evaluar transversalmente las evaluaciones realizadas a los programas de Sedesol en el ejercicio fiscal 2002, privilegiando la descripción y el análisis metodológico general, con la finalidad de conocer el nivel de cumplimiento del conjunto de indicadores temáticos diseñados para la revisión de sus métodos, procesos y resultados, por encima del estudio detallado y en profundidad de cada evaluación aplicada. Lo anterior se complementa brevemente con el estudio longitudinal de la evolución de algunas evaluaciones seleccionadas.

Para el llenado de los indicadores elegidos se empleó información documental, fundamentalmente los informes de evaluación de cada programa, que se complementó con entrevistas semiestructuradas a los diversos actores clave vinculados al proceso.

Los indicadores aplicados relacionan el grado en que cada ítem fue considerado y la calidad de su análisis en el total de evaluaciones realizadas (31). Los mismos podrían presentarse en forma de porcentajes pero, dado que el tamaño del universo es bastante reducido, se consideró mejor hacer sólo referencia al número de casos.

## Métodos y procesos de las evaluaciones externas

### *Fuentes de información y análisis de datos*

Como se revisó en el capítulo 9, el artículo cuarto del Acuerdo no establece lineamientos sobre este aspecto, limitándose a señalar los resultados esperados de la evaluación, entre los que se incluye la verificación en campo del cumplimiento de metas físicas y financieras del programa, la que implícitamente indica que habrá de recogerse información directa, de tipo primario.

Los términos de referencia señalan, en forma muy general, los métodos de recolección y tratamiento de datos, que en todos los casos revisados incluyen trabajo de gabinete, con base en la información secundaria disponible, y trabajo empírico, mismo que debería permitir el análisis costo-beneficio y costo-eficacia, que casi no se intentan en la práctica.

Se trata entonces de analizar si los métodos señalados tienen la amplitud y profundidad suficientes para brindar los datos necesarios a una evaluación de calidad.

Casi la totalidad de las 31 evaluaciones revisadas incluyeron, entre sus *fuentes de información*, documentos y bases de datos estadísticos, y aplicaron encuestas y entrevistas, con pleno cumplimiento de los lineamientos normativos revisados.

Respecto de la información secundaria entregada por las autoridades de los programas, en 13 casos los evaluadores manifiestan diversos problemas: insuficiencia, poca confiabilidad, mala calidad y falta de oportunidad.

En el caso de las entrevistas, éstas se dirigen a informantes-clave (funcionarios y, en menor medida, miembros de las comunidades de los beneficiarios), mediante una selección especialmente dirigida a quienes disponen de información privilegiada.

Por el contrario, en el caso de las encuestas con pretensiones de representatividad (como sucede en la mayoría de la evaluaciones), éstas requieren de la correcta determinación del tamaño muestral y la composición del mismo, de manera que todos los individuos del universo tengan igual probabilidad de ser seleccionados mediante la aplicación sistemática de algún método de muestreo (aleatorio simple, aleatorio estratificado, por conglomerados, etcétera).

En el año en estudio, Sedesol proporcionó a los evaluadores el tamaño muestral. Con este mecanismo Sedesol buscaba que la muestra se determinara de la manera más objetiva y transparente posible, que no la hicieran los operadores ni los evaluadores. Sin embargo, se suscitaron problemas debido a que la empresa contratada no conocía suficientemente los programas.

Este procedimiento, más allá de su costo, que la Sedesol estimó no importante, provocó retrasos significativos en las actividades de planeación y realización del trabajo de campo; pero, sobre todo, dejó una gran insatisfacción en los evaluadores, que no estuvieron de acuerdo con su determinación y solicitaron cambios o propusieron muestras alternativas.<sup>75</sup>

Cabe señalar que, si bien en determinadas condiciones no se toma en cuenta el tamaño del universo para la determinación del correspondiente a la muestra, el listado del mismo facilita la composición aleatoria de ésta; sin embargo, dicha información no se encuentra disponible en la mayoría de los casos.<sup>76</sup>

Esta situación dificulta el análisis de representatividad de la muestra, que se agrava debido a la ausencia casi total de exposición del método empleado en su determinación.

Como puede observarse en el cuadro 8, sólo siete de los 31 programas analizados disponen de la cuantificación de su población objetivo, misma que, en varios casos, coincide aproximadamente con la estimación realizada por Sedesol de 19'225,833 habitantes en las 263 regiones prioritarias definidas, lo que muestra que no ha sido calculada específicamente para el programa en cuestión. El tamaño muestral se presenta en forma completa en 21 casos (aunque sólo en un par de casos con información del método que permitió su determinación). Finalmente, sólo en dos es posible conocer el porcentaje de beneficiarios que fue encuestado y así contar con una idea aproximada de su nivel de representatividad (véase gráfica 8). Éste resulta sumamente bajo (0.002 y 0.006 por ciento), o sea, entre dos y seis encuestados por cada 100,000 personas.<sup>77</sup>

Si se toman en cuenta las reflexiones de Weiss (1999) al respecto, se observa el poco respeto que se tiene al lector cuando no se informa sobre las fuentes de información básicas en que se apoya el informe presentado o simplemente se afirma que la muestra es representativa, sin fundamentarlo.

<sup>75</sup> Un ejemplo lo constituyó el programa de Promoción de las Culturas Indígenas, que incluye entre sus beneficiarios aproximadamente a 12'000,000 de indígenas y una población, no indígena sin estimar, en que la "muestra representativa" aportada por Sedesol fue de 400 beneficiarios de uno solo de los siete subprogramas incluidos.

<sup>76</sup> Nuevamente esto puede ejemplificarse en los programas del INI, en que, si bien la población indígena en estado de pobreza está cuantificada en forma aproximada como acabamos de mencionar, hay programas que se dirigen sólo a una parte no estimada de esta población (ejemplo: jóvenes en edad de estudiar), y más aún, programas que incluyen a población no indígena, interesada en la información sobre dicho grupo social (ejemplo: programa de investigación, bibliotecas, mapotecas, hemerotecas, etcétera). Esta situación se repite en otros programas como Investigación para el Desarrollo Local o Capacitación para el Fortalecimiento de la Sociedad Civil, ambos de Indesol.

<sup>77</sup> Las encuestas de Progreso incluían una persona de cada 100 familias.

CUADRO 8  
TAMAÑOS DE LOS UNIVERSOS Y SUS MUESTRAS

<i>Programas</i>	<i>Tamaño de la población objetivo</i>	<i>Tamaño de la muestra aplicada</i>	<i>Porcentaje muestral población objetivo</i>
1. Desarrollo Humano Oportunidades	20'000,000*	S/muestra	No corresp.
2. Abasto Social de Leche	ND	ND	ND
3. Tortilla	ND	500 beneficiarios	ND
		250 testigos	
4. Abasto Rural	29'152,000	960 beneficiarios	0.005
		960 testigos	
5. Promoción de las Culturas Indígenas	ND	409	ND
6. Desarrollo Económico Productivo Pueblos Indígenas	ND	ND	ND
7. Planeación	ND	495	ND
8. Bienestar Social	ND	367	ND
9. Reparación y Equipamiento Albergues	1,081 albergue	Pers. 94 albergue	8.7 nivel albergue
10. Operación de Albergues	1,081 albergue	185 empleados	ND
		489 beneficiarios	
		66 autoridades comunitarias	
11. Promoción y Procuración de Justicia	ND	264	ND
12. Vinculación y Transferencia de Tecnología Plan Acuario y Programa Desarrollo Habitantes Semidesierto (PDHS)	ND	450 Plan Acuario 500 PDHS	ND
13. Compras, Créditos, Concursos y Asistencia Técnica	ND	350**	ND
14. Expertos en Acción	ND	672 beneficiarios 304 testigos.***	ND
15. Coinversión Social	ND	75 OSC	ND
16. Investigación para el Desarrollo Local	ND	Responsables 36 proyectos	ND
17. Capacitación Fortalecimiento Sociedad Civil	ND	ND	ND
18. Fortalecimiento Institucional. Gestión Social Municipal	ND	No se solicitó	No corresponde
19. Ahorro y Subsidios Vivienda Progresiva	ND	315 viviendas****	ND
20. Identidad Jurídica	20'000,000	S/muestra	No corresponde
21. Zonas de Alta Marginación e Indígenas	ND	374	ND
22. Microrregiones	20'000,000 y 6'600,000 foc.	376	0.002
23. Atención a Jornaleros Agrícolas	ND	601	0.006

<i>Programas</i>	<i>Tamaño de la población objetivo</i>	<i>Tamaño de la muestra aplicada</i>	<i>Porcentaje muestral población objetivo</i>
24. Desarrollo Pueblos y Comunidades Indígenas	ND	S/dato beneficiarios 34 testigos	ND
25. Empleo Temporal	ND	ND	ND
26. Estatales por Demanda	ND	117 miembros Comités	ND
27. Iniciativa Ciudadana 3x1	ND	80 responsables proyectos	ND
28. Jóvenes por México	ND	405	ND
29. Mujeres Jefas de Familia	ND	342	ND
30. Oportunidades Productivas	6'275,677	328 beneficiarios 99 testigos	ND
31. Superación Pobreza Urbana	ND	697	ND

Fuente: Elaboración propia a partir de informes finales de evaluación, 2003, salvo especificación en contrario.

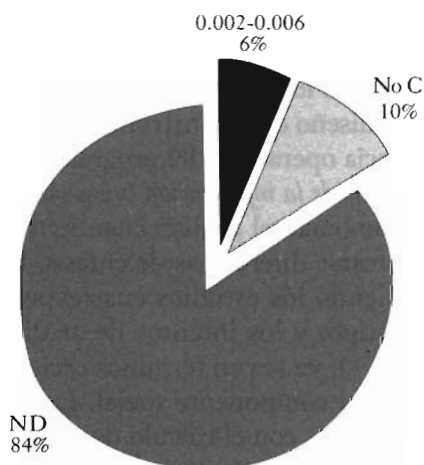
Notas: \* Como el dato no es mencionado en la evaluación, se toma el que brinda la SHCP, que corresponde a la estimación de habitantes en pobreza extrema, ubicados en las 263 microrregiones prioritarias.

\*\*El tamaño diseñado fue de 1,021 beneficiarios, pero sólo se localizaron 350.

\*\*\*El tamaño original era de 1,020 beneficiarios y 264 testigos.

\*\*\*\*El tamaño originalmente determinado fue de 338 viviendas.

GRÁFICA 8  
TAMAÑO MUESTRAL



Por otra parte, únicamente en nueve casos se menciona una encuesta a no beneficiarios que funciona como grupo de control (diseño cuasiexperimental), para lograr mayor precisión en la medición de impactos.

Otro asunto vinculado a la confiabilidad de los resultados de las encuestas aplicadas es la claridad y pertinencia del cuestionario diseñado para ese objetivo, mismo que se presenta en muy pocos casos,<sup>78</sup> además de la forma en que el mismo se aplica, que no es posible verificar en el presente trabajo.<sup>79</sup>

Concluyendo sobre este aspecto, puede afirmarse que el fundamento metodológico de lo que se reporta en los informes, con base en la aplicación de encuestas, resulta de una debilidad muy grande. Si los datos provienen de poblaciones sin significancia estadística, las categorías de análisis que los utilizan carecen de sentido.

En forma complementaria, se aplican otros instrumentos de recolección de información, entre los que se privilegia la entrevista, que resulta aplicada en 20 casos. La misma se realiza por vía telefónica en dos casos, se califica de semiestructurada y en profundidad en seis casos, y de grupal en uno. Se percibe una tendencia a calificar de “en profundidad” a entrevistas que dejan muchas dudas al respecto; también se destacan las debilidades de la aplicación de una encuesta telefónica, en que no es posible confirmar la identidad del locutor ni observar formas de lenguaje no verbal, entre otras.

Cabe anotar también que, en algunas evaluaciones, existe confusión en el uso del término “entrevista”, que aparece como sinónimo de “encuesta”.

Finalmente, se mencionan otros métodos de recolección de información: observación no participante en cuatro casos, dos estudios de casos, empleo de cédulas, monitoreo de precios, censos y encuestas nacionales, grupos focales y observación participante, en un caso cada uno.

Sorprende que en uno de los trabajos, el evaluador, quien debe gozar de toda independencia en el diseño de sus instrumentos, manifieste que sometió a aprobación de la instancia operadora del programa su guión de entrevista.

En cuanto al *tratamiento de la información* (véase cuadro 9), prácticamente todos las evaluaciones combinan el análisis cuantitativo y el cualitativo, por lo que sólo pueden plantearse diferencias de énfasis. Pocos utilizan técnicas especializadas, sobresaliendo los estudios cuasiexperimentales con grupo tratamiento y grupo testigo, y los intentos de analizar el costo/beneficio (C/B) y el costo/eficacia (C/E), ya sea en términos exclusivamente económico-financieros o incluyendo el componente social. La mayoría de las evaluaciones confunde estos estudios con el cálculo de costos unitarios. En menor grado se emplean métodos de estadística inferencial y econometría, análisis financieros (cálculo de periodos de recuperación de la inversión, TIR, VAN,

<sup>78</sup>En algunos casos se menciona que se presenta en un anexo que no se encuentra integrado al informe.

<sup>79</sup>Algunos funcionarios entrevistados reconocieron que el trabajo de campo realizado tuvo fallas importantes (ejemplo: Indesol).



CUADRO 9  
TRATAMIENTO DE LA INFORMACIÓN

<i>Programas</i>	<i>Cuasiexperimentación</i>	<i>C/B o C/E</i>	<i>Estadística Inf. y Econ.</i>	<i>Análisis financiero</i>	<i>Énfasis cualitativo</i>	<i>Indicadores</i>
1. Desarrollo Humano Oportunidades	X		X		X	
2. Abasto Social de Leche	X	X	X	X		
3. Tortilla	X	X	X	X		
4. Abasto Rural	X		X			
5. Promoción de las Culturas Indígenas						X
6. Desarrollo Económico Productivo Pueblos Indígenas						X
7. Planeación						
8. Bienestar Social						
9. Reparación y Equipamiento Albergues						
10. Operación de Albergues						
11. Promoción y Procuración de Justicia					X	
12. Vinculación y Transferencia de Tecnología, Plan Acuario y Programa Desarrollo Habitantes Semidesierto		X				X
13. Compras, Créditos, Concursos y Asistencia Técnica						
14. Expertos en Acción	X	X		X		
15. Coinversión Social						
16. Investigación para el Desarrollo Local						
17. Capacitación Fortalecimiento Sociedad Civil						
18. Fortalecimiento Institucional para la Gestión Social Municipal						
19. Ahorro y Subsidios para la Vivienda Progresiva						
20. Identidad Jurídica					X	
21. Zonas de Alta Marginación e Indígenas						
22. Microrregiones						
23. Atención a Jornaleros Agrícolas						
24. Desarrollo Pueblos y Comunidades Indígenas						
25. Empleo Temporal	X				X	
26. Estatales por Demanda						
27. Iniciativa Ciudadana 3x1	X				X	
28. Jóvenes por México		X				
29. Mujeres Jefas de Familia						
30. Oportunidades Productivas	X	X				
31. Superación Pobreza Urbana	X	X				
Total de casos	9	7	4	3	5	3

Fuente: Elaboración propia a partir de informes finales de evaluación, 2003.

etcétera), técnicas con un marcado énfasis cualitativo, o se diseñan indicadores de evaluación.

En el caso de Oportunidades, es importante mencionar que, al aumentar la población cubierta en 2000-2002, el diseño evaluativo que se venía aplicando se debilitó (parte de su grupo de control pasó a formar parte del grupo de tratamiento), por lo que en 2002, se sustituyó la elección aleatoria de los grupos (que desde 1997 venía usándose en Progresá) por la utilización de la técnica de emparejamiento. Aún no existe ningún documento oficial que analice las consecuencias de esta modificación.

Otro problema de este programa es su insuficiente medición de impacto, realizada en 2002 a nivel de escuelas y clínicas de salud, por lo que en el siguiente año se planeó la aplicación de encuestas individuales.

A nivel de la Dirección General de Evaluación de Sedesol prevalece la idea de que, en la medida de lo posible, es deseable la aplicación del método experimental a la evaluación de todos los programas a cargo de Sedesol, ya que se le considera riguroso y científico. Sin embargo, se reconoce que es muy complejo deslindar resultados entre los programas y los demás factores intervinientes, por lo que se está trabajando en la integración de un padrón único y se cuenta con un proyecto en desarrollo para evaluar el conjunto de políticas sociales. También que, en materia social, no todo se puede hacer con técnicas econométricas y que el análisis cualitativo ofrece buenos sustentos cuando hay insuficiencia de información cuantitativa.

En general, se puede concluir que existe un tratamiento aceptable de la información; sólo se presentan tres casos en que el trabajo es fuertemente descriptivo, limitándose a presentar cuadros y gráficas que ordenan los datos, en lugar de ser más analítico. Son también muy pocos los que intentan aplicar algunos métodos más específicos, los que en varios casos no resultan claros o definitivamente presentan problemas en la selección de la información utilizada en el cálculo.

En resumen, existe un énfasis en la aplicación de entrevistas y encuestas de representatividad pocas veces fundamentada, que se complementa con un tratamiento medianamente aceptable de los datos recogidos.

### *Pertinencia y diseño del programa*

Ni el acuerdo ni los términos de referencia solicitaron el análisis de ninguno de los elementos de este apartado en 2002, lo que equivale a suponer que los programas estaban bien diseñados, por lo que sólo se justifica evaluar sus

resultados e impactos.<sup>80</sup> Ya hace mucho tiempo que Dror (1977) señaló el peligro de aplicar bien políticas y programas mal diseñados, razón por la que se considera parte imprescindible de una evaluación, el considerar si las necesidades y demandas sociales justifican la existencia del programa actual, y de ser así, si hay coherencia entre sus objetivos y dichas necesidades, así como si es viable alcanzar los objetivos propuestos mediante la estrategias diseñadas.

Legisladores entrevistados estuvieron de acuerdo en la necesidad de su inclusión como temas a evaluar, ya que, en la medida en que esto no es solicitado a los evaluadores, sus trabajos muestran pocas referencias al respecto.

Como puede observarse en el cuadro 10, sobre la existencia comprobada de necesidades que justifiquen la existencia del programa sólo hay 14 casos que analizan el tamaño de la población afectada y la magnitud de los problemas que la aquejan. En algunos informes (seis) sólo se describen las características de la población beneficiada o la congruencia del programa con el PND y el PNDS (nueve).

Correspondería luego analizar si el programa incluye un buen diseño de objetivos, pero por lo general, los informes se limitan a enunciarlos. A veces, su definición resulta tan general e indiscutible como “contribuir al logro de mejores condiciones de vida para los beneficiarios” o “aumentar las oportunidades de empleo e ingreso”, por lo que, probablemente, los evaluadores no consideran relevante su análisis. En estos casos resultaría conveniente avanzar hacia el estudio de objetivos operativos o incluso de metas cuantificables.

Luego sería necesario analizar la congruencia entre las estrategias aplicadas y los objetivos previos. En el cuadro 11 se presentan los aspectos considerados en los informes revisados.

Finalmente, sería relevante analizar si el diseño de las reglas de operación respeta la filosofía, los objetivos y estrategias del programa.

En resumen, hay referencias concretas a esta parte del diseño del programa en sólo ocho informes, el análisis de coherencia entre objetivos y necesidades se ve limitado por la insuficiente identificación de las últimas y la congruencia entre las estrategias aplicadas y los objetivos previos sólo es considerada explícitamente por la mitad de los informes.

Se concluye que se hace muy poca evaluación del diseño de los programas, salvo en el aspecto de su focalización, que si bien sólo es solicitado explícitamente sólo por dos programas, constituye la principal estrategia aplicada a los programas de superación de la pobreza.<sup>81</sup>

<sup>80</sup> Para los años posteriores a este análisis (2003 y 2004) se ubicaron términos de referencia que solicitan el análisis del diseño. Consultar, por ejemplo, los relativos al programa Iniciativa Ciudadana 3×1, 2003, p. 19 o el programa Adultos Mayores en Zonas Rurales, 2004, pp. 23 a 25.

<sup>81</sup> Ésta no se aplica en algunos programas como Investigación para el Desarrollo Local y parte del contenido de Promoción de las Cultura Indígenas.

CUADRO 10  
PERTINENCIA DEL PROGRAMA

<i>Programas</i>	<i>Población afectada y magnitud problema</i>	<i>Características población beneficiaria</i>	<i>Congruencia con PND y PNDS</i>
1. Desarrollo Humano Oportunidades	X		
2. Abasto Social de Leche	X		
3. Tortilla	X		
4. Abasto Rural	X		
5. Promoción de las Culturas Indígenas	X	X	
6. Desarrollo Económico Productivo Pueblos Indígenas			
7. Planeación			
8. Bienestar Social	X		
9. Reparación y Equipamiento Albergues	X		
10. Operación de Albergues	X		
11. Promoción y Procuración de Justicia			
12. Vinculación y Transferencia de Tecnología, Plan Acuario y Programa Desarrollo Habitantes Semidesierto		X	X
13. Compras, Créditos, Concursos y Asistencia Técnica		X	X
14. Expertos en Acción			
15. Coinversión Social		X	
16. Investigación para el Desarrollo Local			X
17. Capacitación Fortalecimiento Sociedad Civil			X
18. Fortalecimiento Institucional Gestión Social Municipal	X		
19. Ahorro y Subsidios Vivienda Progresiva	X		X
20. Identidad Jurídica	X		
21. Zonas de Alta Marginación e Indígenas	X		
22. Microrregiones	X		
23. Atención a Jornaleros Agrícolas			
24. Desarrollo Pueblos y Comunidades Indígenas			X
25. Empleo Temporal			X
26. Estatales por Demanda			
27. Iniciativa Ciudadana 3x1	X	X	
28. Jóvenes por México		X	
29. Mujeres Jefas de Familia			
30. Oportunidades Productivas			X
31. Superación Pobreza Urbana			X
Total de casos	14	6	9

Fuente: Elaboración propia a partir de informes finales de evaluación, 2003.

CUADRO 11  
DISEÑO DEL PROGRAMA

<i>Programas</i>	<i>Diseño de objetivos</i>	<i>Congruencia estrategias/ objetivos</i>
1. Desarrollo Humano Oportunidades		
2. Abasto Social de Leche		
3. Tortilla		
4. Abasto Rural	X	X
5. Promoción de las Culturas Indígenas	X	X
6. Desarrollo Económico Productivo Pueblos Indígenas		
7. Planeación		X
8. Bienestar Social	X	X
9. Reparación y Equipamiento Albergues		
10. Operación de Albergues		
11. Promoción y Procuración de Justicia	X	
12. Vinculación y Transferencia de Tecnología, Plan Acuario y Programa Desarrollo Habitantes Semidesierto		
13. Compras, Créditos, Concursos y Asistencia Técnica		X
14. Expertos en Acción		
15. Coinversión Social		X
16. Investigación para el Desarrollo Local		X
17. Capacitación Fortalecimiento Sociedad Civil		X
18. Fortalecimiento Institucional Gestión Social Municipal		
19. Ahorro y Subsidios Vivienda Progresiva		X
20. Identidad Jurídica	X	X
21. Zonas de Alta Marginación e Indígenas	X	X
22. Microrregiones		X
23. Atención a Jornaleros Agrícolas		
24. Des. Pueblos y Comunidades Indígenas		
25. Empleo Temporal		
26. Estatales por Demanda	X	
27. Iniciativa Ciudadana 3×1		X
28. Jóvenes por México		
29. Mujeres Jefas de Familia		
30. Oportunidades Productivas	X	X
31. Superación Pobreza Urbana		X
Total de casos	8	15

Fuente: Elaboración propia a partir de informes finales de evaluación, 2003.

En relación con este aspecto, el cuadro 12 muestra que los elementos más considerados son los mecanismos de identificación y recertificación de beneficiarios, aunque a menudo, más que analizar su diseño, sólo se verifica su cumplimiento. En 13 casos se describen las características de la población atendida para verificar sus niveles de pobreza; sin embargo, esto no resulta suficiente porque no garantiza que los atendidos sean los más pobres o que toda la población objetivo esté siendo atendida. Algo muy similar sucede con los casos en que sólo se verifica la pertenencia a zonas de alta y muy alta marginación. En una tercera parte de los informes se intenta, con mayor o menor precisión, calcular el índice de focalización y los errores de inclusión y exclusión.

### *Recursos utilizados*

Los costos y gastos de operación y el cumplimiento de metas financieras son solicitados en el acuerdo y retomados en los términos de referencia, por lo que en todos los informes de evaluación se encuentra alguna referencia total o parcial al tema.

Un análisis completo del punto debería recuperar la propuesta de Cohen (2001), revisada en el primer capítulo, haciendo alusión al gasto total programado, y al ejercido, su evolución en los últimos años y su distribución por tipo de apoyo, estados, etcétera, así como el monto promedio efectivamente transferido por beneficiario y población objetivo y la eficiencia de su aplicación. Sin embargo, no todos estos elementos se solicitan, por lo que muchos informes incluyen información parcial al respecto.

Su revisión nos muestra información muy variada (véase cuadro 13), siendo la más frecuente la referida al presupuesto ejercido, que de todas formas muestra siete omisiones totales. Este indicador, si bien no se solicita expresamente, resulta fundamental para comparar los costos con la valoración de los beneficios sociales, que el acuerdo incluye como el análisis costo-beneficio de los recursos asignados al programa.

En general se revisa la meta financiera, pero pocos informes profundizan en la información sobre la misma, el presupuesto programado o el nivel de subejercicio alcanzado. Sólo cinco evaluaciones estudian la evolución del gasto ejercido, que será muy importante en su comparación con el comportamiento histórico de la cobertura y otros resultados, así como para diseñar escenarios prospectivos. Sorprende que únicamente una de las evaluaciones que considera el tema plantee explícitamente que trabaja a precios constantes, ya que esto demuestra un manejo elemental de la información económica.

CUADRO 12  
FOCALIZACIÓN

<i>Programas</i>	<i>Identificación y recertificación de beneficiarios</i>	<i>Errores de inclusión y exclusión</i>	<i>Caracterización población beneficiaria</i>	<i>Pertenencia de beneficiarios a microrregiones</i>
1. Desarrollo Humano Oportunidades	X	X		
2. Abasto Social de Leche	X	X		
3. Tortilla		X		
4. Abasto Rural		X	X	
5. Promoción de las Culturas Indígenas	X		X	
6. Desarrollo Económico Productivo Pueblos Indígenas			X	
7. Planeación	X			
8. Bienestar Social		X		
9. Reparación y Equipamiento Albergues	X			
10. Operación de Albergues	X			
11. Promoción y Procuración de Justicia				X
12. Vinculación y Transferencia de Tecnología, Plan Acuario y Programa Desarrollo Habitantes Semidesierto	X			X
13. Compras, Créditos, Concursos y Asistencia Técnica	X		X	
14. Expertos en Acción			X	
15. Coinversión Social			X	
16. Investigación para el Desarrollo Local				X
17. Capacitación Fortalecimiento Sociedad Civil	X			X
18. Fortalecimiento Institucional para la Gestión Social Municipal				X
19. Ahorro y Subsidios para la Vivienda Progresiva		X	X	
20. Identidad Jurídica	X			
21. Zonas de Alta Marginación e Indígenas				X
22. Microrregiones		X		X
23. Atención a Jornaleros Agrícolas	X			
24. Desarrollo Pueblos y Comunidades Indígenas	X		X	
25. Empleo Temporal			X	
26. Estatales por Demanda	X			
27. Iniciativa Ciudadana 3x1				X
28. Jóvenes por México		X		
29. Mujeres Jefas de Familia	X		X	
30. Oportunidades Productivas		X	X	
31. Superación Pobreza Urbana	X	X	X	
Total de casos	15	10	13	7

Fuente: Elaboración propia a partir de informes finales de evaluación, 2003.

La mitad de las evaluaciones hacen un buen análisis de la forma en que el presupuesto se distribuye, fundamentalmente por entidades federativas y vertientes del programa, calculando el gasto promedio por apoyo brindado (proyectos, becas, etcétera). Sólo una cuarta parte calcula el gasto promedio por beneficiario del programa, lo que puede explicarse, en muchos casos, por la dificultad para determinar el número de los últimos.

También casi en la mitad de los casos se plantea que los recursos no provienen exclusivamente del gobierno federal, sino que perciben aportaciones de otros ámbitos de gobierno, de los propios beneficiarios y otros actores.

Un pequeño número de casos se refiere al porcentaje de gastos destinados y utilizados en la operación del programa; pero, mientras algunos los identifican con los gastos administrativos o indirectos; otros lo hacen con el gasto sustantivo, efectivamente recibido por los beneficiarios. En realidad no hay una clara diferenciación de los gastos fijos y variables de cada programa; mucho menos se toman en cuenta los gastos efectuados por los beneficiarios para acceder a los servicios o transferencias del programa.

Otro aspecto que llama la atención se relaciona con la presencia de programas que transfieren recursos a otras instancias,<sup>82</sup> sin que se analice su excepcionalidad y el riesgo de que se duplique la contabilidad de los mismos; así como programas que asignan recursos para la evaluación de otros.<sup>83</sup>

Finalmente, habría que considerar que, en aquellos casos en que los programas son manejados con una mezcla de criterios sociales y comerciales, parece necesario complementar este análisis con un estudio financiero más profundo. El mismo se solicita y realiza en los casos de los programas de Liconsa y Diconsa, pero resulta definitivamente insuficiente en los desarrollados por Fonart.

En el cumplimiento de este apartado hay que señalar la presencia de errores conceptuales en el manejo del criterio de eficiencia, que relaciona los resultados alcanzados con los recursos empleados para hacerlo, pues se percibe que algunos evaluadores consideran que basta con calcular los costos para cubrir un análisis que debe interrelacionar dos variables.

Hay que reconocer que el propio acuerdo y los términos de referencia inducen un poco a este error, al solicitar que se analicen aisladamente el cumplimiento de metas físicas y financieras. Si bien la primeras, aisladas de sus costos, nos permiten una medición de eficacia, que tiene sentido

<sup>82</sup> INI, por ejemplo, en el caso del programa de Desarrollo de los Pueblos y Comunidades Indígenas del Ramo 20 de Sedesol.

<sup>83</sup> Caso de Investigación para el Desarrollo Local, que apoyó la evaluación de procesos de capacitación de Oportunidades.



en la evaluación; el mero uso de recursos, sin relación con sus productos, carece de valor,<sup>84</sup> lo que importa es medir la eficiencia con que se aplicaron, y esto requiere relacionar los costos con los resultados alcanzados, lo que no se realiza.

En menor grado que los insumos financieros, pero de importancia relevante para el funcionamiento del programa, es el resto de los recursos disponibles, que tampoco es solicitado a los evaluadores y se incluye en pocos casos. Ellos incluyen a los tradicionales recursos humanos y materiales (instalaciones, muebles, equipos, etcétera), así como los menos tradicionales, vinculados a la información y la capacidad gerencial. Respecto de la información, ya se comentaron las insuficiencias reportadas al hablar de las fuentes; también es raro encontrar alguna referencia a la capacidad para gestionar los programas.<sup>85</sup>

En lo que atañe a los recursos humanos, encontramos menciones a su falta de capacitación, disponibilidad de docentes, número de trabajadores y su evolución, viáticos, remuneraciones y otros incentivos al personal.

En atención a los recursos materiales, se analiza el abasto de medicamentos, la infraestructura escolar, la red informática, la disponibilidad de alimentos, materiales de construcción, vehículos, teleaulas, etcétera. Se trata de pocas menciones que dependen de manera importante del tipo de actividad desarrollada por el programa.

### *Análisis de procesos*

El acuerdo y los términos de referencia tampoco mencionan explícitamente este análisis. De no considerarlo, muchos procesos desarrollados quedarían ocultos en una “caja negra”, desconociéndose aspectos como la productividad, calidad y eficiencia con que los insumos son transformados en productos o servicios.

A pesar de la ausencia de una solicitud expresa, todos los evaluadores incluyen referencias a dichos procesos, aunque con amplitud muy diversa, y cumplen en revisar su apego a las reglas de operación.

En su revisión detallada destacan las menciones a los procesos de difusión y promoción del programa y sus convocatorias, recepción de recursos, requisitos y formalidades del proceso de solicitud y dictamen; selección y recertificación de proyectos y beneficiarios, altas y bajas de los padrones, control y

<sup>84</sup>Un programa que gasta el 100 por ciento de sus recursos no es, sólo por eso, mejor que otro que gastó el 80 por ciento.

<sup>85</sup>A excepción del programa de Operación de Albergues, por ejemplo.

CUADRO 13  
RECURSOS

<i>Programas</i>	<i>Presup. programado</i>	<i>Presup. ejercido</i>	<i>Evolución presupuesto ejercido</i>	<i>Costo promedio por apoyo</i>	<i>Distrib. del gasto</i>	<i>Gasto per capita</i>	<i>Otros financia mientos</i>	<i>Gasto Oper. %</i>	<i>Otros</i>
1. Desarrollo Humano Oportunidades				X					
2. Abasto Social de Leche						X			
3. Tortilla									
4. Abasto Rural		X							Relaciones otros gastos
5. Promoción de las Culturas Indígenas		X		X		X			
6. Desarrollo Económico Productivo Pueblos Indígenas				X					
7. Planeación		X			X	X			
8. Bienestar Social		X		X		X	X		
9. Reparación y Equipamiento Albergues	X			X			X		
10. Operación de Albergues		X	X				X		
11. Promoción y Procuración de Justicia	X	X		X	X	X			
12. Vinculación y Transferencia, de Tecnología, Plan Acuario y Programa Desarrollo Habitantes Semidesierto		X		X	X				
13. Compras, Créditos, Concursos y Asistencia Técnica	X	X	X			X			

14. Expertos en Acción		X		X					
15. Coinversión Social		X			X	X			
16. Investigación para el Desarrollo Local		X	X		X		X		
17. Capacitación Fortalecimiento Sociedad Civil		X	X	X	X		X		
18. Fortalecimiento Institucional para la Gestión Social Municipal	X	X		X	X		X		
19. Ahorro y Subsidios para la Vivienda Progresiva	X	X		X	X				Costos futuros beneficiarios
20. Identidad Jurídica									Costos futuros
21. Zonas de Alta Marginación e Indígenas	X	X		X	X		X		
22. Microrregiones	X	X		X	X		X	X	
23. Atención a Jornaleros Agrícolas		X	X		X	X	X		
24. Desarrollo Pueblos y Comunidades Indígenas	X	X		X	X	X	X	X	
25. Empleo Temporal									
26. Estatales por Demanda		X		X	X	X	X	X	
27. Iniciativa Ciudadana 3×1	X	X		X			X		
28. Jóvenes por México		X		X	X			X	
29. Mujeres Jefas de Familia		X		X	X	X	X	X	
30. Oportunidades Productivas	X	X		X	X			X	
31. Superación Pobreza Urbana		X		X	X		X		
Total de casos	10	24	5	20	16	11	14	6	3

Fuente: Elaboración propia a partir de informes finales de evaluación, 2003.

CUADRO 14  
PROCESOS

<i>Programas</i>	<i>Difusión y promoción</i>	<i>Recepción recursos</i>	<i>Solicitud y dictamen</i>	<i>Selección y recertificación</i>	<i>Control e información</i>	<i>Coordinación</i>
1. Desarrollo Humano Oportunidades				X		
2. Abasto Social de Leche				X		
3. Tortilla				X		
4. Abasto Rural						
5. Promoción de las Culturas Indígenas				X	X	
6. Desarrollo Económico Productivo Pueblos Indígenas			X			
7. Planeación						
8. Bienestar Social		X				
9. Reparación y Equipamiento Albergues		X			X	
10. Operación de Albergues		X				X
11. Promoción y Procuración de Justicia		X			X	X
12. Vinculación y Transferencia de Tecnología, Plan Acuario y Programa · Desarrollo Habitantes Semidesierto						
13. Compras, Créditos, Concursos y Asistencia Técnica	X					
14. Expertos en Acción		X				
15. Coinversión Social			X		X	X
16. Investigación para el Desarrollo Local		X	X			
17. Capacitación Fortalecimiento Sociedad Civil						
18. Fortalecimiento Institucional para la Gestión Social Municipal						
19. Ahorro y Subsidios para la Vivienda Progresiva	X			X		
20. Identidad Jurídica						
21. Zonas de Alta Marginación e Indígenas						
22. Microrregiones						X
23. Atención a Jornaleros Agrícolas		X				
24. Desarrollo Pueblos y Comunidades Indígenas		X	X			X
25. Empleo Temporal		X				X
26. Estatales por Demanda	X	X	X		X	X
27. Iniciativa Ciudadana 3x1	X	X	X			
28. Jóvenes por México	X		X		X	X
29. Mujeres Jefas de Familia	X	X	X			X
30. Oportunidades Productivas		X	X		X	X
31. Superación Pobreza Urbana	X		X		X	
Total de casos	7	13	10	5	8	10

Fuente: Elaboración propia a partir de informes finales de evaluación, 2003.

generación de información institucional, y coordinación con otras instancias (véase cuadro 14). Adicionalmente, se encuentran analizados otros procesos a los que se atribuye menor importancia (ejemplo: calidad y trato adecuado en el proceso educativo de Oportunidades), o sólo son relevantes para algunos programas (ejemplo apertura y cierre de tiendas en Abasto Rural).

A pesar del número limitado y diverso de procesos que se analizan en cada evaluación, puede considerarse que se realiza una revisión aceptable, incluyendo la discusión de los principales problemas identificados. Este análisis podría homogeneizarse, para facilitar comparaciones, mediante un guión de aspectos a considerar, sin perder de vista las particularidades de cada programa evaluado.

Conceptualmente, este análisis vuelve a revelar confusiones entre veracidad y eficacia, eficiencia y tiempo de demora, o eficacia y gasto.

En cuanto al apego de los procesos a las reglas de operación y su transparencia, existe mucha dificultad para su verificación, lo que sólo puede hacerse en función de las percepciones de la muestra seleccionada o mediante el estudio de expedientes. Con base en los últimos, es posible revisar aleatoriamente los resultados de dichos procesos, pero no los procesos mismos en que se tomaron las decisiones. En muchos casos, la información disponible es insuficiente y, prácticamente, se requeriría haber actuado como observador en los procesos para poder fundamentar plenamente un juicio al respecto. Resulta más factible la verificación entre los beneficiarios encuestados de los requisitos de elegibilidad exigidos.

### *Medición de cobertura*

El análisis de la cobertura, su variación respecto del año anterior y su desglose por tipo de apoyo otorgado, por entidad federativa y por núcleo de población o tipo de empresa, son solicitados por el acuerdo.

Se trata de una expectativa bastante ambiciosa, que requiere disponer de buenas fuentes de información institucional, ya que el evaluador sólo puede verificar la cobertura en los lugares seleccionados para la realización de trabajo de campo. Si el programa no dispone del dato, el evaluador no puede obtenerlo porque éste requiere de un trabajo censal. Es por ello que una funcionaria de uno de los programas evaluados manifestó que, en materia de cobertura, metas físicas y financieras, etcétera, el evaluador reproduce las cifras que el programa le aporta, por lo que a la instancia que lo opera no le agrega nada, salvo algunas sugerencias relativas a su sistema de información, personal adicional, etcétera.

CUADRO 15  
COBERTURA

<i>Programas</i>	<i>Población atendida</i>	<i>Porcentaje de población objetivo</i>	<i>Distribución por estados, municipios o localidades</i>	<i>Cálculo exclusivo a nivel estados, municipios o localidades</i>	<i>Evolución de la cobertura</i>	<i>Otros</i>
1. Desarrollo Humano Oportunidades	X				X	
2. Abasto Social de Leche			X		X	Informa el total de población no cubierta
3. Tortilla			X		X	
4. Abasto Rural	X		X			
5. Promoción de las Culturas Indígenas				X		
6. Desarrollo Económico Productivo Pueblos Indígenas				X		
7. Planeación	X					
8. Bienestar Social	X					
9. Reparación y Equipamiento Albergues	X					
10. Operación de Albergues				X		
11. Promoción y Procuración de Justicia	X		X			Distribución por proyectos
12. Vinculación y Transferencia de Tecnología, Plan Acuario y Programa Desarrollo Habitantes Semidesierto	X				X	
13. Compras, Créditos, Concursos y Asistencia Técnica						

14. Expertos en Acción	X				X	
15. Coinversión Social	X					
16. Investigación para el Desarrollo Local			X			
17. Capacitación Fortalecimiento Sociedad Civil					X	
18. Fortalecimiento Institucional para la Gestión Social Municipal					X	
19. Ahorro y Subsidios para la Vivienda Progresiva					X	
20. Identidad Jurídica						
21. Zonas de Alta Marginación e Indígenas		X	X			
22. Microrregiones	X	X				Indica porcentaje de cobertura respecto de la meta
23. Atención a Jornaleros Agrícolas	X					
24. Desarrollo Pueblos y Comunidades Indígenas	X		X			
25. Empleo Temporal	X	X				
26. Estatales por Demanda		X			X	
27. Iniciativa Ciudadana 3x1	X		X			Analiza la pertinencia de los beneficiarios a la población objetivo.
28. Jóvenes por México						
29. Mujeres Jefas de Familia						
30. Oportunidades Productivas	X	X	X			
31. Superación Pobreza Urbana					X	
Total de casos	15	5	9		8	5
						4

Fuente: Elaboración propia a partir de informes finales de evaluación, 2003.

En la mitad de las evaluaciones (véase cuadro 15), se incluye la cantidad de población atendida como cálculo de la cobertura; aun cuando lo más importante es su porcentaje (que sólo cinco evaluaciones calculan), que relaciona la población anterior con la población objetivo total, siempre y cuando la primera forme parte de la última, ya que se presentan casos en que se atiende a personas con características diferentes a las definidas por el programa. La falta de medición de dicho porcentaje puede deberse a que, como ya se mencionó, es muy común que el dato relativo al tamaño de población objetivo no esté disponible, lo que impide contar con información suficiente para valorar el esfuerzo que aún resta por hacer. Otro motivo posible es que el acuerdo y los términos de referencia no especifican el tipo de medición de cobertura esperado.<sup>86</sup>

En los 15 casos en que se revisa la población atendida, existe la posibilidad de que ésta coincida sólo parcialmente con la población objetivo, lo que sólo una de las evaluaciones revisa.

Cerca de la cuarta parte de las evaluaciones analiza la distribución de la cobertura por edad, género, nivel de marginación, etcétera y la estudia a nivel estatal o municipal; sólo en cinco casos se analiza su evolución.

En resumen, 27 de las 31 evaluaciones hacen alguna medición de cobertura, el resto argumenta no contar con información suficiente, la confunde con el logro de metas físicas, no la considera porque el programa aún no se ha aplicado o se limita a discutir la cuantificación hecha por Sedesol.

De esta forma, puede afirmarse que el nivel de cumplimiento de los informes en relación con lo solicitado en el acuerdo es bastante bajo, lo que en algunos casos, es responsabilidad del evaluador, pero en muchos se debe a la insuficiente información institucional.

### *Resultados alcanzados*

El acuerdo y los términos de referencia solicitan el resultado de la evaluación en campo del cumplimiento de las metas físicas; sin embargo, la definición de éstas resulta tan limitada y parcial que todos los informes incluyen otros elementos adicionales.

En relación con dicho cumplimiento, sucede lo mismo que se revisó en el caso de la cobertura: la información que se presenta es institucional, no hay manera de que el evaluador haga su propio cálculo porque esto requeriría de una revisión objetiva y completa en todo el país. Es así que ocho de las

<sup>86</sup> La misma fue aclarada posteriormente en la guía propuesta por la SEP.



CUADRO 16  
RESULTADOS

<i>Programas</i>	<i>Servicios brindados</i>	<i>Logro de objetivos</i>	<i>Distribución regional de acciones</i>	<i>Evolución de número de acciones</i>	<i>Distribución por género</i>
1. Desarrollo Humano Oportunidades	X			X	
2. Abasto Social de Leche	X	X			
3. Tortilla	X	X			
4. Abasto Rural	X	X		X	
5. Promoción de las Culturas Indígenas	X	X			
6. Desarrollo Económico Productivo Pueblos Indígenas	X				
7. Planeación	X	X		X	X
8. Bienestar Social	X	X			
9. Reparación y Equipamiento Albergues	X	X			
10. Operación de Albergues	X	X		X	X
11. Promoción y Procuración de Justicia	X	X	X		
12. Vinculación y Transferencia de Tecnología, Plan Acuario y Programa Desarrollo Habitantes Semidesierto	X	X			
13. Compras, Créditos, Concursos y Asistencia Técnica	X	X	X		
14. Expertos en Acción	X	X	X	X	X
15. Coinversión Social	X				
16. Investigación para el Desarrollo Local	X	X		X	
17. Capacitación Fortalecimiento Sociedad Civil	X	X			X
18. Fortalecimiento Institucional para la Gestión Social Municipal	X	X			
19. Ahorro y Subsidios para la Vivienda Progresiva	X	X	X		
20. Identidad Jurídica	X	X			
21. Zonas de Alta Marginación e Indígenas	X	X			
22. Microrregiones	X	X			
23. Atención a Jornaleros Agrícolas	X				
24. Desarrollo Pueblos y Comunidades Indígenas	X	X	X		X
25. Empleo Temporal					
26. Estatales por Demanda	X	X	X		
27. Iniciativa Ciudadana 3x1	X	X	X		
28. Jóvenes por México	X	X	X	X	X
29. Mujeres Jefas de Familia		X			
30. Oportunidades Productivas	X	X			
31. Superación Pobreza Urbana					
Total de casos	28	25	8	7	6

Fuente: Elaboración propia a partir de informes finales de evaluación, 2003.

evaluaciones efectuadas fundamentan la imposibilidad de la comprobación exhaustiva de acciones en todo el país, las verifican parcialmente en campo y a través de las percepciones de funcionarios y beneficiarios.<sup>87</sup>

El análisis de los informes de evaluación muestra (véase cuadro 16) en este apartado: una descripción del número de servicios brindados, en casi todos los casos; un análisis de la eficacia alcanzada en el logro de objetivos (no así de su eficiencia), en una amplia mayoría de los programas evaluados; la distribución de las acciones realizadas por regiones o estados, la evolución del número de acciones realizadas, y el análisis de la equidad de género en aproximadamente la cuarta parte de las evaluaciones realizadas.

En dos casos, se realiza también un análisis crítico de los indicadores propuestos por Sedesol y/o su dificultad para alimentarlos con información.<sup>88</sup>

En general, se procede a un tratamiento suficiente de los servicios brindados y el cumplimiento de metas, con pocas excepciones. Son muy pocos los que profundizan en el análisis de su evolución y distribución. Un buen número de casos analiza críticamente tanto la información que se solicita como los indicadores propuestos para su sistematización. Excepcionalmente se perciben confusiones conceptuales como la presentación de resultados bajo el título de impactos, o errores en la manera de calcular la eficacia y el costo-efectividad.

### *Impactos provocados*

En términos de impactos, el acuerdo y los términos de referencia solicitan analizar los beneficios económicos y sociales que resultan de la aplicación de los recursos asignados a cada programa; no se refieren al análisis del efecto logrado en la solución del problema encarado por el programa, que corresponde estrictamente al impacto. También requieren el costo-beneficio de los recursos asignados al programa, en función de costos y gastos unitarios en que se incurre para operar el programa.

La descripción de los beneficios económicos y sociales se cubre, con diferente nivel de detalle, en casi todos los programas,<sup>89</sup> si bien algunas evaluaciones los confunden con los resultados (número de acciones, proyectos, becas, etcétera). Pero, lo que provoca dificultades es el intento por identificar

<sup>87</sup> Se trata de Promoción de las Culturas Indígenas, Planeación, Reparación y Equipamiento de Albergues, Bienestar Social, programas de Conaza, Fortalecimiento Institucional para la Gestión Social Municipal, Estatales por Demanda y Mujeres Jefas de Familia.

<sup>88</sup> Estatales por Demanda e Iniciativa Ciudadana.

<sup>89</sup> Con excepción de Investigación para el Desarrollo Local y de Identidad Jurídica, en que este último caso no operó.

y medir realmente los impactos, y el análisis costo-beneficio, que algunos intentan sin realizar una aplicación correcta del método.

En su análisis, encontramos (véase cuadro 17) que más de la cuarta parte de las evaluaciones omiten todo tratamiento de los impactos. En los casos restantes se analizan los cambios en las condiciones socioeconómicas de los beneficiarios (empleo, ingresos reales, educación, vivienda, alimentación, servicios básicos, etcétera), el impacto percibido y las expectativas de los beneficiarios a nivel de las familias, las organizaciones y las comunidades, o se incluye una valoración global del impacto en opinión del evaluador, en nueve programas. Sólo en seis casos se verifican los cambios atribuidos al programa mediante la comparación del grupo de tratamiento y de otro testigo. Adicionalmente, se mencionan referencias específicas (ejemplo: impacto en la migración en los casos de los programas de Zonas de Alta Marginación e Indígenas y Microrregiones).

Por otra parte, también se presentan opiniones que cuestionan la posibilidad de medir impactos por las siguientes razones: falta de información institucional suficiente sobre los beneficios del programa y ausencia de diagnóstico basal, intangibilidad de los impactos como dificultad para su cuantificación, falta de definición de metas en este indicador, y necesidad de disponer de plazos mayores para su medición.

Oportunidades es el único programa que, en 2002, ofreció una cifra de población que superó la pobreza extrema (5 por ciento), muy criticada en el medio académico y periodístico. La Dirección General de Evaluación de Sedesol lo considera el programa más avanzado en materia de medición de impacto y espera que el resto evolucione en ese sentido.

Sin embargo, su diseño a nivel de escuelas (en lugar de un seguimiento de becarios) es muy débil en materia de medición de impacto; sólo se cuantifica la inscripción y retención escolar pero no evalúa el logro educativo.

El programa de Empleo Temporal, si bien analiza cambios en las condiciones de vida de los beneficiarios, no se focaliza al estudio del empleo, sino que le concede la misma importancia que a la educación y demás indicadores de bienestar.

En algunos casos se reportan impactos que no corresponden, por estar logrados con recursos ajenos al programa, o se confunden los beneficios económicos y sociales con datos presupuestales.

Por otra parte, habría que evaluar los aprendizajes generados (desarrollo de la cultura evaluativa, de tecnologías apropiadas, etcétera), que resultan más difusos, o incluso impactos provocados por el solo hecho de que los funcionarios sepan que existirá una evaluación.

CUADRO 17  
IMPACTOS

<i>Programas</i>	<i>Cambios en condiciones de beneficiarios</i>	<i>Atribución de los cambios al programa</i>	<i>Percepciones y expectativas de los beneficiarios</i>	<i>Valoración global del impacto</i>
1. Desarrollo Humano Oportunidades	X	X		
2. Abasto Social de Leche	X	X		
3. Tortilla	X	X		
4. Abasto Rural	X	X		
5. Promoción de las Culturas Indígenas				
6. Desarrollo Económico Productivo Pueblos Indígenas			X	
7. Planeación			X	X
8. Bienestar Social			X	X
9. Reparación y Equipamiento Albergues			X	X
10. Operación de Albergues				
11. Promoción y Procuración de Justicia			X	X
12. Vinculación de Transferencia de Tecnología, Plan Acuario y Programa Desarrollo Habitantes Semidesierto				
13. Compras, Créditos, Concursos y Asistencia Técnica				
14. Expertos en Acción			X	
15. Conversión Social	X			
16. Investigación para el Desarrollo Local			X	
17. Capacitación Fortalecimiento Sociedad Civil			X	
18. Fortalecimiento Institucional para la Gestión Social Municipal				X
19. Ahorro y Subsidios para la Vivienda Progresiva	X			
20. Identidad Jurídica				
21. Zonas de Alta Marginación e Indígenas	X			
22. Microrregiones			X	
23. Atención a Jornaleros Agrícolas	X	X		
24. Desarrollo Pueblos y Comunidades Indígenas				
25. Empleo Temporal	X	X		X
26. Estatales por Demanda				
27. Iniciativa Ciudadana 3x1			X	X
28. Jóvenes por México				X
29. Mujeres Jefas de Familia				
30. Oportunidades Productivas	X			X
31. Superación Pobreza Urbana	X			
Total de casos	11	6	10	9

Fuente: Elaboración propia a partir de informes finales de evaluación, 2003.

En conclusión, se cuenta con un esfuerzo muy variado por realizar algo más que una mera descripción de beneficios económicos y sociales, que las dos terceras partes de las evaluaciones logran mediante mediciones de impacto, con distinto grado de objetividad.<sup>90</sup> La SHCP considera que los indicadores que se utilizan no trascienden el mero seguimiento de datos de entrada y de salida.

Ante esta situación, considerada común a muchos países en opinión del BM, diversos legisladores estuvieron de acuerdo en la necesidad de intentar mediciones de impacto sólo en periodos plurianuales. La ASF, por ejemplo, tarda entre tres y cinco años para volver a medir el impacto de un mismo programa. Lo mismo sucede en algunos casos nacionales revisados en el capítulo 4 (Francia y Canadá).

### *Participación y satisfacción social*

La revisión de la participación social y ciudadana, estrategia privilegiada por el PND, no es solicitada por el acuerdo; los términos de referencia sólo la requieren expresamente en dos programas. Por el contrario, es ampliamente solicitado el análisis del grado de satisfacción de la población objetivo.

Trece evaluaciones no mencionan el tema, mientras muchos informes hacen referencia a la ausencia de participación o a los bajos niveles alcanzados por la misma en la gestión de los programas. Las referencias específicas de participación en la gestión de los programas son pocas (véase cuadro 18): formación de comités de beneficiarios, asambleas, integración de comités de contraloría social, consideración de las opiniones de la comunidad, participación de beneficiarios en cursos de capacitación, promoción de la participación ciudadana a través del involucramiento de OSC, aportaciones en dinero, especie, o trabajo. También se informa sobre la participación obligatoria en la conformación de comités proobras en el caso del Programa para los Habitantes del Semidesierto (PHSD) de Conaza.

Hay que resaltar que los beneficiarios son considerados para la obtención de información para los informes (mediante la aplicación de encuestas, entrevistas, grupos focales, etcétera), pero no existe ninguna referencia a que hayan participado en los procesos mismos de evaluación.<sup>91</sup> La Dirección

<sup>90</sup> Al iniciar 2005, la Dirección General de Evaluación informó que sólo se han realizado nueve evaluaciones específicas de impacto, algunas de ellas vinculadas a los programas vigentes en 2002 en los casos de: Progreso-Oportunidades, Liconsá, Diconsá y la estrategia de Microrregiones.

<sup>91</sup> Sólo las organizaciones apoyadas por Indesol, en algunos casos (Equipo Pueblo, por ejemplo), aplican procesos de evaluación participativa de sus proyectos.

CUADRO 18  
PARTICIPACIÓN

<i>Programas</i>	<i>Organizaciones de beneficiarios</i>	<i>Cursos de capacitación</i>	<i>Organizaciones de la sociedad civil</i>	<i>Aportaciones en dinero o especie</i>	<i>Aportaciones de trabajo</i>
1. Desarrollo Humano Oportunidades					
2. Abasto Social de Leche	X				
3. Tortilla	X				
4. Abasto Rural	X	X			
5. Promoción de las Culturas Indígenas	X				
6. Desarrollo Económico Productivo Pueblos Indígenas	X				
7. Planeación	X	X			
8. Bienestar Social	X				
9. Reparación y Equipamiento Albergues					X
10. Operación de Albergues					X
11. Promoción y Procuración de Justicia					
12. Vinculación y Transferencia de Tecnología, Plan Acuario y Programa Desarrollo Habitantes Semidesierto					
13. Compras, Créditos, Concursos y Asistencia Técnica					
14. Expertos en Acción					
15. Coinversión Social				X	X
16. Investigación para el Desarrollo Local				X	
17. Capacitación Fortalecimiento Sociedad Civil			X		
18. Fortalecimiento Institucional para la Gestión Social Municipal					
19. Ahorro y Subsidios para la Vivienda Progresiva					
20. Identidad Jurídica					
21. Zonas de Alta Marginación e Indígenas		X			
22. Microrregiones					
23. Atención a Jornaleros Agrícolas					X
24. Desarrollo Pueblos y Comunidades Indígenas					
25. Empleo Temporal					
26. Estatales por Demanda		X			
27. Iniciativa Ciudadana 3x1		X			
28. Jóvenes por México					
29. Mujeres Jefas de Familia				X	
30. Oportunidades Productivas					
31. Superación Pobreza Urbana				X	
Total de casos	10	3	4	2	2

Fuente: Elaboración propia a partir de informes finales de evaluación, 2003.

General de Evaluación confirmó que sólo participan mediante la presentación de quejas o en respuesta a las encuestas aplicadas.<sup>92</sup>

A nivel de los evaluadores, SHCP, SFP y OSC hubo amplio acuerdo sobre el bajo nivel de participación social, con sólo dos excepciones de organizaciones que lo calificaron como medio y de la evaluación cualitativa aplicada en el caso de Oportunidades.

Entre los miembros del Poder Legislativo, los entrevistados afirman que no conocen ninguna evaluación realizada con participación social, o al menos, en que ésta haya jugado un papel relevante. Algún congresista de la oposición llegó a manifestar que “ésta no le gusta al gobierno ni sabe cómo encauzarla”, por lo que el borrador de reglamento de la LGDS reduce mucho su peso original, limitándose a la creación de consejos consultivos de carácter honorario.

Ante tal ausencia de participación de los beneficiarios en la toma de decisiones y los procedimientos de evaluación, surgió la inquietud de investigar si alguna organización de la sociedad civil (también promovida en el discurso de los planes y programas gubernamentales) participaba en los procesos de evaluación gubernamental. Para ello se estableció contacto con unas 15 OSC y consultoras especializadas en ellas, de las cuales seis reportaron tener experiencia evaluativa,<sup>93</sup> por lo que se le aplicó una entrevista a las cinco que fue posible contactar.

Lo anterior permitió conocer que ninguna de las entrevistadas había tenido relación alguna con los procesos de evaluación externa de programas gubernamentales en estudio; sin embargo, en uno de los casos la organización había sido contratada para realizar una evaluación al programa Organización Productiva de Mujeres Indígenas para la Comisión Nacional de Desarrollo de los Pueblos Indígenas (INI en 2002), y en otro, realizó una evaluación independiente al Progreso, dentro de la campaña “El Banco Mundial en la Mira de las Mujeres” (Red Nacional de Asesoras y Promotoras Rurales).

En el resto de los casos, la experiencia evaluativa se vinculaba con evaluaciones hechas en consultoría para los programas de otras OSC nacionales, a nivel internacional, o referidas a la evaluación de sus propios proyectos o programas.

Es de destacar, por contraste con las evaluaciones externas a programas gubernamentales, el cuidado que las OSC ponen en los mecanismos participa-

<sup>92</sup> Su director manifiesta promover la participación de los *stakeholders* involucrados, pero hace referencia sólo a los funcionarios de su dirección, los operadores y evaluadores (Hernández, 2005).

<sup>93</sup> Se contactó a SIPAM, CIDHAL, FUNDAR, etcétera, y reportaron contar con experiencias en evaluación: CAM; Kinal Antsetil; GEM, Espiral Consultora S.C., Red Nacional de Asesores y Promotores Rurales, y una persona que perteneció a Cidhal y trabaja actualmente como consultora independiente.

tivos y en la discusión y difusión de los resultados de la evaluación a través de talleres, espacios de diálogo, publicaciones, etcétera, que forman parte del aprendizaje organizacional. En general, los resultados son utilizados para conocer avances y efectuar ajustes a la planeación por la misma organización o para informar a la agencia de cooperación.

Estas experiencias, consideran las entrevistadas, se traducen en aprendizajes de contenido (aclarar el rumbo), de metodología (desarrollo de instrumentos técnicos y formas de motivar a participar), de conocimiento del contexto nacional, funcionamiento del gobierno, etcétera, que complementan la mirada interna de los implicados en el programa.

Frente a este breve diagnóstico sobre la participación social en los procesos evaluativos, uno de los evaluadores externos entrevistados concluyó que es necesario elevar y potenciar a los sujetos de la evaluación, construir evaluación y no aplicarla de fuera y mecánicamente como algo dado.

En la práctica se trata de una elección ideológica de fondo, como manifestó un congresista entrevistado: una visión tecnocrática de transferencias de efectivo, en forma individualizada, para ser utilizadas en el mercado; o una organización social de los beneficiarios participando en los procesos decisivos como, en su opinión, habría intentado hacer el Pronasol.

En cuanto a la satisfacción de los beneficiarios, ésta es recogida en la mayoría de las evaluaciones mediante la aplicación de encuestas; sin embargo, y a pesar de ser solicitada de manera explícita, hay 11 programas que no la revisan.

Lo anterior muestra poco apego de los evaluadores a un aspecto solicitado, importante y fácil de cubrir, así como un poco de mayor atención al tema de la participación, también muy importante, pero no solicitado.

### *Escenarios prospectivos*

El requerimiento de su presentación es señalado en el acuerdo y recogido en diversos términos de referencia; sin embargo, son muy pocas las evaluaciones que los intentan:

- cuatro evaluaciones incluyen el diseño de un escenario tendencial y otro alternativo;
- tres informes organizan otros tipos de escenarios,
- cuatro presentan tendencias o previsiones mínimas, sin construir escenarios, y
- tres presentan argumentaciones, no muy convincentes, de por qué no se incluyeron.



Como puede percibirse, éste constituye el aspecto menos cubierto de todos los solicitados en el acuerdo y hace suponer que, o bien los evaluadores no tuvieron tiempo suficiente para realizarlo, ya que es el último paso del proceso solicitado; o bien hay muy poca *expertise* desarrollada en el tema.

### *Valoración global del programa*

Como se recordará de la lectura del capítulo 2, toda evaluación supone, para ser considerada como tal, la elaboración final de un juicio valorativo, sistemáticamente fundamentado, sobre la gestión del programa y sus logros. El Acuerdo solicita al respecto: opinión sobre la permanencia o conclusión del programa, conclusiones y recomendaciones.

Como puede verse en el cuadro 19, 12 trabajos cumplen plenamente con este juicio valorativo explícito y lo acompañan de recomendaciones para su continuación, uno evalúa ampliamente el programa sin presentar un juicio sintético, y 17 casos se limitan a realizar recomendaciones que, implícita o explícitamente, suponen la continuidad del programa. En los dos restantes no se incluye ni juicio valorativo ni recomendaciones.

En conclusión, siendo rigurosos, sólo un poco más de la tercera parte de los trabajos cumplen con este requisito para ser verdaderas evaluaciones y existen tres casos insuficientes en materia de conclusiones.

### *Utilización de resultados y recomendaciones de evaluaciones previas*

Este aspecto no fue incluido en el acuerdo; sin embargo, fue solicitado, por lo menos, en los términos del referencial de un programa, y permite dar continuidad a los trabajos realizados, así como detectar la utilidad práctica de las recomendaciones efectuadas.

Se recuerda que siete de los programas revisados comenzaron a funcionar en 2002,<sup>94</sup> y otros dos parecen no haber sido evaluados previamente, aunque esto no se menciona. En consecuencia, en estos casos, el punto carece de sentido.

De los 22 programas restantes, su cumplimiento se verifica sólo en 12 informes. Lo anterior deja 10 programas que, a pesar de tener antecedentes

<sup>94</sup>Identidad Jurídica, Desarrollo de Pueblos y Comunidades Indígenas, Estatales por Demanda, Iniciativa Ciudadana, Mujeres Jefas de Familia, Oportunidades Productivas y Superación de la Pobreza Urbana.

CUADRO 19

## VALORACIÓN DEL PROGRAMA Y RECOMENDACIONES

<i>Programas</i>	<i>Juicio y recomendaciones explícitas</i>	<i>Recomendaciones con juicio implícito</i>
1. Desarrollo Humano Oportunidades		
2. Abasto Social de Leche		X
3. Tortilla		X
4. Abasto Rural	X	
5. Promoción de las Culturas Indígenas	X	
6. Desarrollo Económico Productivo Pueblos Indígenas	X	
7. Planeación	X	
8. Bienestar Social		
9. Reparación y Equipamiento Albergues	X	
10. Operación de Albergues	X	
11. Promoción y Procuración de Justicia	X (no sintético)	
12. Vinculación y Transferencia de Tecnología, Plan Acuario y Programa Desarrollo Habitantes Semidesierto	X	
13. Compras, Créditos, Concursos y Asistencia Técnica		X
14. Expertos en Acción		X
15. Coinversión Social	X	
16. Investigación para el Desarrollo Local		X
17. Capacitación Fortalecimiento Sociedad Civil		X
18. Fortalecimiento Institucional, Gestión Social Municipal		X
19. Ahorro y Subsidios Vivienda Progresiva.	X	
20. Identidad Jurídica		X
21. Zonas de Alta Marginación e Indígenas		X
22. Microrregiones		X
23. Atención a Jornaleros Agrícolas		X
24. Desarrollo Pueblos y Comunidades Indígenas		X
25. Empleo Temporal		
26. Estatales por Demanda		X
27. Iniciativa Ciudadana 3×1	X	
28. Jóvenes por México	X	
29. Mujeres Jefas de Familia		X
30. Oportunidades Productivas		X
31. Superación Pobreza Urbana		X
Total	12	16

Fuente: Elaboración propia a partir de informes finales de evaluación, 2003.

de evaluación previa, no los refieren, presentando su evaluación como un esfuerzo aislado.

### *Informes finales*

En este apartado se refiere el contenido y claridad de los informes de evaluación y sus resúmenes ejecutivos. En teoría se espera que un buen informe explicita sus objetivos, describa el contexto en que se produjo la evaluación, la metodología aplicada y los motivos de su elección, sus resultados y criterios de valoración, así como que presente un juicio sintético del desempeño del programa y realice recomendaciones pertinentes; pero, el análisis de los informes finales de evaluación muestra una cobertura muy limitada de estos temas.

El funcionario entrevistado en la SFP manifestó que el principal problema en el cumplimiento de sus funciones de revisión de los informes lo constituye el desorden, ya que la información no se presenta organizada de acuerdo a los temas estipulados en el acuerdo. Agregó que se mencionan conclusiones por todos lados, se incurre en problemas metodológicos y se identifican incongruencias entre resultados, conclusiones y recomendaciones.

La Dirección General de Evaluación de Sedesol también considera que los informes deben mejorarse mucho y las recomendaciones sustentarse adecuadamente.

### Costo/calidad de los resultados de las evaluaciones externas

Como se señaló en el capítulo 6, Schwandt y Halpern (1988, citado en Bustello, 2001) sugieren que la metaevaluación tome en cuenta, además de la metodología utilizada, la bondad de los resultados hallados en la evaluación. Como la última puede resultar fuertemente impactada por los recursos financieros disponibles para su realización, mismos que en nuestro contexto presentan una gran variación en sus montos, se considera necesario tomar en cuenta ambas variables en forma combinada. Se propone entonces una clasificación del conjunto de las evaluaciones, con base en el presupuesto destinado al trabajo y la calidad de su producto.

Esta última surge como consecuencia del comportamiento del conjunto de indicadores revisados. Para su medición agregada, Schwandt y Halpern recomiendan tomar en cuenta la coherencia lógica existente entre los cuatro elementos identificados por Patton: datos, interpretaciones, valoraciones y

recomendaciones. En nuestro caso, y con todas las limitaciones que implica utilizar sólo información seleccionada, se retomó uno de los métodos de medición de calidad sugeridos por Abate (2000), muy cercano a la propuesta de Patton, y adaptado a nuestro contexto: una combinación de tres de los elementos revisados en esta metaevaluación, considerados parámetros significativos para su medición: 1. contar con un diseño, recolección y tratamiento de datos rigurosos; 2. cumplir con los requerimientos especificados en el acuerdo y los términos de referencia; y 3. concluir con la explicitación de un juicio valorativo fundamentado sobre el programa.

Como resultado de lo anterior se obtienen: 1. siete evaluaciones de alta calidad, cuatro de las cuales coinciden con la valoración efectuada por la Dirección General de Evaluación (Oportunidades, Abasto Social de Leche, Tortilla y Abasto Rural); 2. siete de calidad media; y 3. las 16 restantes con calidad baja<sup>95</sup> (véase cuadro 20). Ninguna de las evaluaciones consideradas cumplió plenamente con todos los parámetros considerados, razón por la que ninguna alcanzó un nivel que pudiera calificarse como óptimo.

Los resultados anteriores fueron reforzados a través de otro de los cuatro métodos identificados por Abate: la percepción promedio que sobre la calidad tienen los diferentes actores del proceso, misma que fue recogida en las entrevistas. Una amplia mayoría las consideró “regulares”; en particular, cuatro de los cinco evaluadores entrevistados fueron de esta opinión aunque, al igual que los legisladores, reconocieron una gran variabilidad. El funcionario entrevistado en la SFP las calificó contundentemente como “malas”.

La variedad de calidades identificada coincide también con la opinión manifestada por el director general adjunto de Gestión Institucional, quien mencionó la existencia de un abanico de calidades en las evaluaciones realizadas: con problemas de información, falta de datos estandarizados, de carácter más descriptivo que explicativo, poco útiles para el consumo interno, etcétera, por lo que consideró necesaria la realización de trabajos de metaevaluación.

El resultado del estudio combinado del monto de recursos empleados y la calidad de evaluaciones muestra la falta de correlación entre estas variables: la mitad de las evaluaciones exhiben congruencia entre los recursos aplicados y la calidad lograda, cinco presentan una calidad superior a sus costos y 10 la situación inversa (véase cuadro 20).

El nivel más alto de presupuesto corresponde ampliamente al programa de Desarrollo Humano Oportunidades; sin embargo, a pesar de su alta calidad, no es posible juzgar a su evaluación como óptima, debido a que no

<sup>95</sup> Una no se tomó en cuenta por tratarse de un programa casi sin actividad, cuya evaluación es muy acotada.

CUADRO 20  
 CLASIFICACIÓN DEL NÚMERO DE EVALUACIONES  
 SEGÚN CALIDAD Y COSTO

	<i>Recursos bajos (menos de 400,000 pesos)</i>	<i>Recursos medios (entre 400,000 y 1'000,000 de pesos)</i>	<i>Recursos altos (entre 1'000,000 y 5'000,000 de pesos)</i>	<i>Recursos muy altos (más de 5'000,000 de pesos)</i>	<i>Total</i>
Calidad baja	9	5	2	0	16
Calidad media	1	4	2	0	7
Calidad alta	0	4	2	1	7
Calidad muy alta	0	0	0	0	0
Total	10	13	6	1	30

Fuente: Elaboración propia a partir de cuadros previos.

cumple con algunos de los requisitos establecidos en el acuerdo y no integra los diversos trabajos realizados, ni emite un juicio valorativo de conjunto. Esto se fundamenta también en que el diseño de grupos de tratamiento y control a nivel de escuelas y centros de salud no resulta suficiente para medir impactos, en que se debería trabajar a nivel de los beneficiarios. Además, Oportunidades tendría que medir la sinergia que aporta la reunión de servicios de salud, alimentación y educación en lugar de la realización de tres programas separados.

Abasto Rural resulta la evaluación más cara después de Oportunidades y, de acuerdo con los parámetros establecidos, es el que más se acerca a una calidad óptima, conjuntamente con las evaluaciones efectuadas a los programas de Conaza e Iniciativa Ciudadana 3×1.

En el caso de Abasto Social de Leche, los recursos, siendo los más altos después de Oportunidades y Abasto Rural, distan mucho de los anteriores; sin embargo, se obtiene un trabajo de buena calidad, en el que se lamenta la ausencia de un juicio valorativo final.

En el extremo opuesto, resulta preocupante que dos evaluaciones de alto costo logren sólo un nivel bajo de calidad.

El análisis realizado permite concluir que se cuenta con un limitado avance de la evaluación de programas de desarrollo social en México. También verificamos que esto es así en la percepción de prácticamente todos los

actores entrevistados en torno al proceso,<sup>96</sup> por lo que se les preguntó sobre las causas que consideran provocan esta situación.

Los resultados de sus respuestas, que se resumen en el cuadro 21, indican que las más importantes son de índole cultural, social y política, seguidas de las económicas y administrativas. Sorprende que contando con tantos problemas metodológicos e insuficiencias de formación en la materia, las razones de ese tipo sean las menos nombradas. Esto sugiere falta de conciencia sobre las limitaciones que las actuales evaluaciones presentan, mismas que han quedado demostradas en el análisis de los informes correspondientes.

### Juicio valorativo de la metaevaluación transversal

A las limitaciones financieras y de recursos humanos especializados, identificadas en el capítulo anterior, el actual agrega problemas en los procedimientos, el manejo metodológico y la calidad de los resultados hallados.<sup>97</sup> Entre los primeros se señalan imprecisiones y ausencias por parte de la demanda (acuerdo, bases y términos de referencia).

En relación con el manejo metodológico, salvo pocas excepciones, se identifican evaluaciones con base en la aplicación de instrumentos a muestras de representatividad desconocida o dudosa, se cubren parcialmente los aspectos solicitados y otros de reconocida relevancia para arribar a conclusiones bien fundamentadas, y se emiten pocos juicios explícitos de valoración del programa analizado.

Respecto del tercer problema, se destaca la baja calidad identificada en 16 de las evaluaciones analizadas y la falta de correlación con sus respectivos costos, así como una conciencia insuficiente de los entrevistados de que los problemas metodológicos constituyen una causa importante de la situación anterior.

Se insiste en que el juicio expresado no se dirige exclusivamente a la calidad del trabajo de los evaluadores, sino a la alcanzada por el informe final presentado, misma que depende también de factores vinculados a la organización y gestión del sistema de evaluación en su conjunto, algunos de los cuales son de responsabilidad gubernamental.

<sup>96</sup> Sólo estuvieron en desacuerdo dos OSC en relación con sus propias evaluaciones y la ASF que considera como evaluaciones a los ensayos, diagnósticos y estudios de prospectiva en torno a la situación económica, social y cultural del país.

<sup>97</sup> Díaz y Rosales (2003: 322) concluyen en forma similar sobre la limitada calidad de las evaluaciones que revisan en el contexto costarricense y señalan que "...lo más sintomático es que los evaluadores evidencian desconocimiento de las teorías y las técnicas propias de la disciplina...".

CUADRO 21

### RAZONES DEL LIMITADO AVANCE DE LA EVALUACIÓN EN MÉXICO (Número de consideraciones)

<i>Actores entrevistados (número de respuestas)</i>	<i>Razones culturales</i>	<i>Razones económicas</i>	<i>Razones sociales</i>	<i>Razones políticas</i>	<i>Razones administrativas</i>	<i>Razones metodológicas</i>	<i>Razones psicológicas</i>	<i>Otras razones</i>
Funcionarios Sedesol(1)					1	1		
Funcionarios programas (4)	3	3	3	2	4	1	2	Faltan cuadros para coordinación de evaluación
Evaluadores externos (5)	4	2	3	3	4	1	1	Falta cultura de la crítica y democracia
SHCP (1)	1	1	1	1				Ley de Adquisiciones no garantiza calidad
SFP (1)	1	1	1	1				
IFAI (1)			1	1				
OSC (5)	3	1	3	3		3	2	
Congreso (3)	3		1	1				Técnicas
Total	15	8	13	12	9	6	5	
Núm. total de respuestas recibidas: 17								

Fuente: Elaboración propia a partir de la información recogida en entrevistas.

Nota: Sólo dos OSC consideraron que la evaluación no era limitada.

En ese sentido, salvo un evaluador que declaró no tener ningún problema en su desarrollo, los demás entrevistados coincidieron ampliamente en torno a las principales dificultades detectadas en el entorno de su proceso de trabajo:

- la falta de una cultura de evaluación que provoca resistencias en las instancias responsables de los programas;
- la rigidez de la normatividad, que prácticamente induce la metodología y resultados de la evaluación, impidiendo la creatividad y la incorporación de nuevos elementos metodológicos al análisis;
- el presupuesto insuficiente para realizar un trabajo de campo a fondo;

- la falta de suficiente capital humano formado en evaluación para redactar términos de referencia, calificar propuestas técnicas y asignar contratos;
- la ausencia de información confiable, ordenada, completa y oportuna acerca de la operación de los programas a evaluar;
- el tiempo asignado, que obliga a estar más pendiente de las entregas administrativas que de la reflexión en profundidad; y
- las dificultades para conciliar el trabajo como evaluador e investigador.

### Evolución de algunas evaluaciones seleccionadas

Ante la imposibilidad de analizar la evolución de las 31 evaluaciones revisadas en el año 2002, se seleccionaron tres de características bien diferenciadas como aproximación a la realización de un estudio longitudinal (véase cuadro 22):

-*Desarrollo Humano Oportunidades*: es un programa que cuenta con una evaluación diseñada desde su inicio y el mayor número de evaluaciones realizadas (considerando también al programa antecedente), con amplia información y una evaluación en 2002 de muy alto costo y buena calidad, mas no excelente, de acuerdo con los indicadores seleccionados y el criterio aplicado en el capítulo previo.

-*Microrregiones*: se trata de un programa creado en 2002, uno de los pocos que cuenta con algún nivel de participación social y que fue evaluado con recursos medios, logrando un nivel de calidad congruente con éstos.

-*Promoción de Culturas Indígenas*: es un programa sectorizado muy antiguo y desarticulado, dirigido a la población que sufre los mayores niveles de pobreza,<sup>98</sup> con muy poca información y uno de los más reducidos presupuestos para su evaluación, que mediante la aplicación de técnicas tradicionales y el diseño de indicadores obtuvo una calidad media, según el mismo procedimiento aplicado en los casos previos.

Se encontró que las evaluaciones realizadas en 1998-2000 al Progreso mostraban mayor experiencia y coordinación del equipo responsable, así como precisión en el uso del lenguaje; mientras las efectuadas a partir de 2001 sufrieron el debilitamiento del diseño experimental debido a la integración al padrón de beneficiarios del grupo de control, lo que se acompaña de una inapropiada definición de sus unidades de análisis (escuelas, centros de salud o municipios incorporados a Oportunidades).

<sup>98</sup>Según el Banco Mundial (2004a), en 2002, el 44 por ciento de los indígenas se encontraban en el quintil más pobre de ingreso, constituyendo el 20 por ciento de los pobres extremos.



CUADRO 22  
**CARACTERÍSTICAS DE LOS PROGRAMAS SELECCIONADOS  
 Y DE SU EVALUACIÓN**

<i>Programas</i>	<i>Grupo</i>	<i>Presupuesto ejercido programa 2002 (millones de pesos)</i>		<i>Cobertura poblacional programa</i>	<i>Recursos (miles de pesos)/ calidad evaluación (2002)</i>	<i>Énfasis en técnicas de evaluación empleadas</i>	<i>Números de equipos evaluadores en 2001-2003</i>
			<i>Auditado ASF</i>				
Oportunidades	Sectorizado	5,965	Sí (en sector salud)	4'240,000 familias	Recursos muy altos (28,882) / calidad alta	Cuantitativas	1*
Microrregiones	Ramo 20	220.6	Sí	854,000 beneficiarios	Recursos medios (970)/ calidad media	Cualitativas	2
Promoción Culturas Indígenas	Sectorizado (INI)	57.6	No	s/inf. precisa	Recursos bajos (244)/ calidad media	Indicadores cuantitativos y cualitativos	3

Fuente: Informes, entrevistas y datos aportados por el IFAI.

\*Nota: Formalmente el contrato en 2001 fue otorgado al CIDE y en los años restantes al INSP, sin embargo, el equipo se ha compuesto siempre por las mismas instituciones.

En la comparación de sus evaluaciones 2002-2003 destaca la desincorporación en el equipo evaluador de personas con experiencia, la pérdida de referencia de los datos a la gestión evaluada, el menor cumplimiento de la normatividad que regula la contratación del trabajo, y la reducción de la extensión y contenido del informe, manteniendo los problemas metodológicos y las insuficiencias en el tratamiento del tema, ya identificados.

En el caso de Microrregiones (denominado Desarrollo Local en 2003) la misma comparación muestra: tendencia al economicismo con pérdida de profundización en el estudio de los problemas sociales y políticos y las percepciones de los entrevistados, debilitamiento del tratamiento de la focalización, la revisión del cumplimiento de las reglas de operación, y la identificación de impactos.

Por su parte, Promoción de Culturas Indígenas (llamado Desarrollo de Culturas Indígenas en 2003) presenta tres subprogramas que dejan de evaluarse debido a su desincorporación del programa y otros dos que práctica-

CUADRO 23  
EVOLUCIÓN DE LOS COSTOS DE LAS EVALUACIONES  
DE PROGRAMAS, 2002-2003

<i>Programas</i>	<i>Costo de la Evaluación 2002 (miles de pesos)</i>	<i>Costo de la Evaluación 2003 (miles de pesos)</i>	<i>Porcentaje de incremento de costos</i>
1. Desarrollo Humano Oportunidades	28,881.89	46,265.18	60.19
2. Abasto Social de Leche	1,150.00	1,483.85**	28.96
3. Programa Tortilla	498.00	518.71	4.16
4. Abasto Rural	2,080.00	2,503.96***	20.33
5. Compras, Créditos, Concursos y Asistencia Técnica	1,132.81	670.00	-40.86
6. Coinversión Social	150.00*	919.43	-14.14
		por los programas números 6 a 9	
7. Investigación Desarrollo Local	221.00		
8. Capacitación Fortalecimiento de la Sociedad Civil	700.00		
	por los programas números 8 y 9		
9. Fortalecimiento Institucional Gestión Social Municipal			
10. Identidad Jurídica	655.30	1,580.00	-28.68
		por los programas números 10 a 12 integrados	
11. Zonas de Alta Marginación e Indígenas	568.40		
12. Microrregiones	970.01		
13. Atención Jornaleros Agrícolas	848.58	1,495.00	76.17
14. Empleo temporal (incluye emergencias)	2,001.10 SCT 2,926.38 Sagarpa	1,827.70 SCT 979.76 Sagarpa	- 43.03
15. Iniciativa Ciudadana	576.26	967.50	67.89
16. Jóvenes por México	800.00	862.50	7.81
17. Oportunidades Productivas	725.00	1,538.24	112.17
Totales	44,884.73	61,611.83	37.27

Fuente: Elaboración propia a partir de información aportada por el IFAI.

Notas: \* Si bien la cifra que se nos entregó oficialmente es la que se consigna, tanto la responsable de la evaluación como la funcionaria del programa, ambas entrevistadas, manifestaron que el costo fue muy diferente (600,000 pesos), monto que se considera más confiable.

\*\* Se consigna la cifra aportada por el programa; la de Sedesol fue de 1'300,000 pesos.

\*\*\* La Sedesol reportó 2'262,000 pesos.

mente no se incluyen en el informe correspondiente, se debilita la medición de cobertura y otros resultados, y desaparece el análisis de focalización y el diseño de escenarios.

En consecuencia, sin pretender que los casos estudiados sean representativos de todas las evaluaciones realizadas, las tres evaluaciones analizadas en 2003 muestran una peligrosa tendencia a la disminución de la calidad del trabajo efectuado.<sup>99</sup>

Por otra parte, el cuadro 23 muestra la evolución de los costos de los programas cuyos datos se disponen para ambos años (2002 y 2003). Salvo los casos de Oportunidades Productivas, Atención a Jornaleros Agrícolas e Iniciativa Ciudadana 3x1, que presentaron importantes incrementos de costos probablemente debidos a que se integraron con algunos de los 14 programas desaparecidos, ninguna de las evaluaciones de las que se dispone de información en 2002 y 2003 incrementó sus costos en la forma en que lo hizo Oportunidades, incluso algunas los disminuyeron. Los incrementos alcanzados por Abasto Social de Leche y por Abasto Rural se justifican además por la aplicación de técnicas adicionales de recolección de información (nuevos grupos focales, etcétera).

A partir de 2001 Oportunidades ha utilizado la información de las ENCEL aplicadas por INEGI, por lo que si bien se comprende la extensión y complejidad del programa, sus evaluaciones no parecen justificar los costos pagados por su realización (especialmente la última, cuyo contenido es mucho más limitado que la previa). A partir del momento en que las evaluaciones comienzan a realizarse en el país, los costos se disparan: la evaluación 2003 costó aproximadamente cuatro veces más que las realizadas por el IFPRI en 1999 y 2000, y 60 por ciento más que la de 2002, lo que no se justifica ni por inflación ni por mayor cantidad de trabajo o mejor calidad del mismo. Esto resulta más preocupante debido a la ausencia de procedimientos de licitación pública para designar a las instituciones evaluadoras.

## Comentarios finales

El capítulo que finaliza ha permitido la realización de un análisis estático, transversal, de la situación global de la evaluación de programas sociales en 2002, que identifica un importante avance en el desarrollo de la función evaluativa misma, aunque aún con muchos problemas por resolver: muestreos de representatividad desconocida o muy baja, tratamiento

<sup>99</sup> Por el contrario, el director general de evaluación manifiesta que los informes de los evaluadores han mejorado (Hernández, 2005).

insuficiente de la información, inclusión parcial de aspectos importantes y pocos juicios de valoración global de la gestión del programa. En la relación costo/calidad se identifican 15 evaluaciones con niveles congruentes en estas dos dimensiones, cinco con calidad más alta que sus costos y otras 10 con la situación inversa. Los entrevistados identificaron razones culturales, sociales y políticas como explicación de los problemas descritos.

Los informes analizados, a luz de las normas para la evaluación de programas del Joint Committee on Standards for Educational Evaluation (véase anexo 1), enfrentan serios problemas de precisión en torno a la descripción suficiente, validez, confiabilidad e imparcialidad de la información (P4, P5, P6 y P11) y en participación de los agentes involucrados (U1); dudas importantes en el manejo de conflictos de intereses (C7), la justificación de las conclusiones (P10) y relación costo/efectividad de las evaluaciones realizadas (F3); y ausencia de juicios de valor o poca fundamentación de los mismos (U4), así como de una valoración completa y justa de las fortalezas y debilidades del programa (C5).

Por su parte, si se procede a valorar dichos informes con base en el Key Evaluation Checklist de Scriven (1991), se vuelven a detectar como problemas relevantes algunos de los ya identificados en el caso anterior: la insuficiencia de información sobre los costos monetarios y no monetarios incluidos en la evaluación y, principalmente, la falta de explicitación de los criterios de mérito en que se fundamentan los pocos juicios de valoración emitidos y de una síntesis de todo lo analizado que sirva de sustento a los mismos.

En cuanto al análisis de la evolución de las evaluaciones de algunos programas, los mismos muestran situaciones y tendencias muy desiguales en materia de costos, que impactan el tipo de evaluación que es posible realizar. También sugieren una tendencia a disminuir la calidad de estas evaluaciones en el tiempo.

Cabe, entonces, preguntarse por qué, a medida que los equipos ganan en experiencia, no se está logrando mejorar la calidad. Además de los recursos financieros disponibles, habría que estudiar, por un lado, el grado de desarrollo del capital humano alcanzado en la materia; en segundo lugar, los mecanismos de selección entre las propuestas recibidas en las convocatorias para ver hasta qué punto garantizan la calidad del proyecto ganador; y por último, la generación de incentivos suficientes por parte del sistema de evaluación organizado para que dicho equipo emplee al máximo el conocimiento adquirido.

En torno al primero, parece indiscutible que se ha realizado un aprendizaje, aunque éste podría no ser tan significativo como sería de desear ya que: 1. no existen mayores recomendaciones fundamentadas por parte de la dirección de evaluación de Sedesol, la SHCP, la SFP, la Cámara de Diputados, o reuniones de discusión entre evaluadores y con los beneficiarios, que permitan la retroalimentación de los procesos; 2. al interior de los equipos se producen sustituciones de investigadores pertinentes al tema y con amplia experiencia, por otros que carecen de algunos de esos dos requisitos o de ambos; 3. los evaluadores externos desconocen los resultados a que llega la ASF cuando ésta audita el desempeño del programa, mismos que podrían enriquecer su trabajo; y 4. tampoco se cuenta con programas universitarios que contribuyan a esta formación de evaluadores.

Respecto de los mecanismos de selección, existen varios aspectos que podrían estar limitando los avances en calidad: 1. la escasa difusión de convocatorias y el empleo de algunos mecanismos de adjudicación directa que limitan una mayor competencia técnica entre los equipos disponibles; 2. la obligación impuesta por la ley de adquisiciones de contratar al equipo que, cumpliendo *a priori* con los requerimientos, ofrezca el precio más bajo, lo que no permite tomar en consideración expectativas de calidad; y 3. la cuestionada capacidad de las comisiones formadas *ad hoc* para seleccionar entre los candidatos participantes en una licitación.

Finalmente, en cuanto a la falta de incentivos para utilizar en forma intensiva el aprendizaje logrado, las razones podrían ser también diversas:

1. un cierto esclerosamiento de los equipos de trabajo que comienzan a aplicar en forma mecánica los mismos procedimientos a programas diferentes;
2. una tendencia de los mismos a acumular evaluaciones, de manera de realizar economías de escala que se traduzcan en mayor rentabilidad, con el consecuente riesgo de no atenderlas todas en la forma deseable; y
3. la percepción de que los informes son muy poco utilizados y no garantizan la obtención de un nuevo contrato, por lo que esto deja de ser un aliciente para realizar un buen trabajo.

Como consecuencia de los análisis realizados en el presente capítulo y en el precedente, se valora positivamente que se haya iniciado un proceso de evaluación de programas sociales y se considera explicable que subsistan problemas en su implementación. Sin embargo, no se soslaya la cantidad

y magnitud de los problemas identificados, especialmente los organizativos y metodológicos, los que aún requieren que se tomen medidas para contribuir a su mejoramiento a corto y mediano plazo, ya que muchas evaluaciones no se justifican en términos de costo/nefecio. En cambio, forman parte de un proceso de aprendizaje hacia el perfeccionamiento de su calidad intrínseca.

## Capítulo 12

# Impactos y perspectivas de la evaluación de programas de desarrollo social

### Introducción

EN ESTE último capítulo se sintetizan los principales resultados a los que se ha ido arribando en el curso de la segunda parte, de manera a poder presentar propuestas de solución a los problemas identificados.

Para ello se caracterizan las evaluaciones revisadas a la luz del marco teórico-metodológico presentado en la primera parte del libro, se analiza su impacto, y se organizan los principales resultados parciales de los capítulos anteriores según un esquema de análisis FODA (fortalezas, oportunidades, debilidades y amenazas) aplicado al sistema de evaluación en su conjunto. De éste se infiere una lista de problemas que se ordenan y relacionan con base en un diagrama de árbol.

Posteriormente se proponen algunas posibles soluciones de mediano plazo a las causas identificadas en las “raíces” del árbol, y también a otras cuestiones que requieren y permiten una intervención con resultados de corto plazo, identificando las instancias que tendrían que ponerlas en marcha.

A continuación se contrasta el escenario tendencial, que se proyecta a partir de la situación actual sin cambios, con otro alternativo, fruto de la implementación de las sugerencias planteadas.

### Caracterización de las evaluaciones externas analizadas

La caracterización siguiente ha sido realizada con apego a los elementos analizados en la segunda parte y al marco presentado en la primera:

#### Conceptos

Los términos de referencia para la realización de evaluaciones externas presentan variantes en el concepto de evaluación; sin embargo, coinciden en

concebirla como un mecanismo para: 1. cumplir con la obligación establecida por el PEF (y ahora por la LGDS), que les garantiza la obtención de recursos para el siguiente año, y 2. ofrecer información para la toma de decisiones en torno al programa evaluado. En los casos en que se alude a un proceso de aprendizaje, éste es sólo de carácter institucional, por lo que el concepto de evaluación puede caracterizarse como de tipo instrumental, más que preocupado por un proceso de aprendizaje social.

#### Tipo de evaluaciones

Se trata de evaluaciones *ex post*, externas, no participativas, parcialmente endoformativas y de carácter cuantitativo-cualitativo, con énfasis en el primero. Pretenden ser sistemáticas y verdaderas aunque a veces pueden caer en la pseudoevaluación. También intentan cubrir la evaluación del contexto, los insumos, procesos, resultados, logro de metas, impactos y satisfacción de la población, pero en la práctica descuidan especialmente los dos últimos aspectos y diseño, por lo que no constituyen evaluaciones integrales.

#### Actitudes

Se busca desarrollarlas con actitud objetiva, de ahí el énfasis en la experimentación y el uso de métodos cuantitativos que las autoridades responsables del proceso desean promover; sin embargo, se corre el riesgo de que algunas de ellas adapten la información a los prejuicios culturales y marcos cognitivos de los “clientes”. La mayoría presenta un énfasis descriptivo, algunas intentan explicar los resultados hallados y raras veces lo hacen con actitud normativa, o sea, realizando un análisis crítico basado en su propio sistema de valores, al margen de las metas institucionales.

#### Enfoques

Las evaluaciones responden a un enfoque sectorial tradicional, de inspiración positivista, centrado en medición y pruebas; por oposición al pluralista que no se ha desarrollado en nuestro contexto, basado en el constructivismo y el énfasis en el aprendizaje, la participación social, el empleo de la conciliación entre actores en conflicto, y la credibilidad de los procesos. En algunos casos se aplica una perspectiva contingente, que retoma elementos de ambos enfoques. También existe interés en aplicar un enfoque territorial por parte de la Dirección General Adjunta de Análisis Territorial de la Sedesol.



### Destinatarios

Los principales destinatarios de los informes de evaluación son la Sedesol y sus instancias operadoras, el Congreso y, en menor grado, la SHCP y la SFP. Los mismos no se dirigen hacia los beneficiarios o la ciudadanía en general.

### Métodos

Como ya se mencionó, la Dirección General de Evaluación de la Sedesol aspira a realizar evaluaciones experimentales con métodos econométricos de fuerte contenido cuantitativo. Debido a las restricciones imperantes en términos de recursos, a la fecha el programa que más se ha acercado a su realización es Oportunidades. También se usan, en menor grado y con algunos problemas, los análisis de Costo/Beneficio, Costo/Eficacia, las herramientas de la estadística descriptiva e inferencial o el análisis financiero. Entre las técnicas cualitativas de recolección de datos, la encuesta es utilizada casi en todos los casos y su información recibe tratamiento estadístico. También se emplea generalmente la entrevista a informantes-clave y la observación tradicional, dentro de los métodos cualitativos, y en menor medida se utilizan dinámicas de grupo, paneles de discusión y se elaboran estudios de caso.

### Criterios

En concordancia con los métodos aplicados, el criterio más empleado es el de eficacia, seguidos de la oportunidad y la satisfacción de los beneficiarios, pero descuidando parcialmente la efectividad, la equidad, la pertinencia, la congruencia interna, la calidad e incluso la eficiencia.

### Especificidades

Los programas sociales presentan características particulares en el contexto de la evaluación gubernamental a las que no se está poniendo suficiente atención en la práctica. Resalta entre ellas la necesidad de contar con plazos medianos y largos para identificar y medir los impactos de los programas en las condiciones de vida de la población, lo que no se respeta al exigir evaluaciones anuales. También hay tendencia a adjudicar al programa los cambios percibidos en el problema, cuando estos pueden ser provocados en parte por algún otro de los diversos programas aplicados o por las múltiples variables contextuales que inciden en él. Finalmente, tampoco se percibe un tratamiento profundo de los intereses en conflicto de los variados actores vinculados a los programas, ni a la particular importancia que en este caso adquiere la

transparencia y la responsabilidad de los funcionarios en la gestión, así como la participación social.

### Dificultades

En opinión de los diferentes actores entrevistados, las mayores dificultades para el desarrollo de las evaluaciones son, en orden de importancia, de carácter: cultural, social, político, administrativo, económico y metodológico. Si bien sólo las OSC atribuyen a las causas metodológicas tanta importancia como a las tres primeras mencionadas, con base en el análisis realizado se concluye que son los problemas de orden metodológico y organizacional los que están limitando más el avance del proceso.

### Impacto de las evaluaciones

Para que el trabajo de evaluación realizado logre tener impacto en el perfeccionamiento de los programas es necesario que sus resultados se difundan y se tomen en cuenta en los procesos decisorios de las diferentes instancias de los poderes Ejecutivo y Legislativo, así como se analicen y discutan entre los ciudadanos en general.

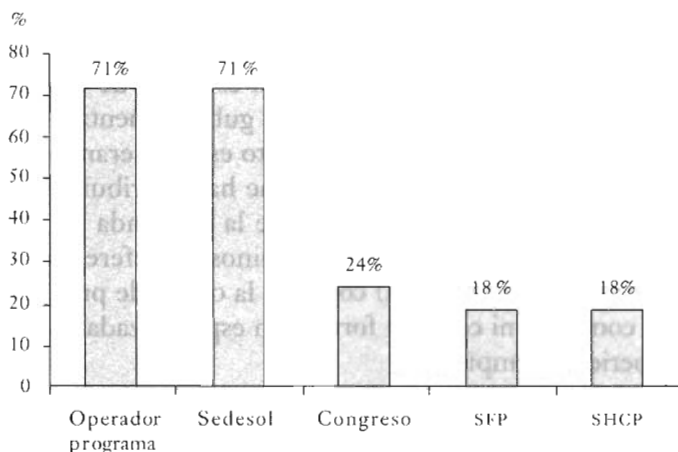
En el caso de las evaluaciones externas los informes correspondientes, o al menos sus resúmenes ejecutivos, han sido publicados en sus páginas web por unos pocos programas (Oportunidades, Abasto Social de Leche, Tortilla y Abasto Rural), mientras la disposición del resto requiere su solicitud por los interesados a través del IFAI. También se cuenta con spots en televisión y un periódico mural en el edificio de Sedesol que reproduce los principales logros. Oportunidades utiliza folletos o posters para informar a los últimos, y realiza seminarios de presentación de resultados, impresión de libros y CD para académicos y funcionarios.

La experiencia de diversos países (en AL principalmente Costa Rica) muestra que no basta realizar una medición de logros de uso interno a la administración pública, sino que su divulgación constituye un medio poderoso de estimulación y control social de la gestión.

En relación con la utilización de los informes en la toma de decisiones, los diferentes entrevistados coinciden en que son la Sedesol y los operadores de programas quienes más utilizan los resultados de las evaluaciones (véase gráfica 9).

Sin embargo, su utilización se percibe más vinculada a la revisión del cumplimiento del convenio de evaluación y a la realización de ajustes a las reglas de operación y términos de referencia, que a las modificaciones sustantivas de los programas.

GRÁFICA 9  
UTILIZACIÓN DE RESULTADOS



Finalmente, se identifica una casi nula utilización por el Congreso, particularmente por parte de la Comisión de Desarrollo Social de la Cámara de Diputados. Su homóloga en el Senado no ha recibido hasta el presente copia de los informes, pero deberá empezar a hacerlo con la entrada en vigencia de la LGDS.

Alamilla (s/f) sintetiza como sigue la situación imperante en esta materia: poca utilización de los resultados por las dependencias y tendencia a cumplir sólo formalmente con las evaluaciones externas; toma de decisiones de la Cámara de Diputados basada en techos financieros preestablecidos y sin tomar en consideración el impacto de los programas, poca transparencia en relación con los ajustes efectuados al proyecto de presupuesto, entrega parcial de los informes finales de las evaluaciones externas de 2002, y revisión general de su calidad por parte de la SFP.

En resumen, hay acuerdo entre las instancias consultadas en el sentido de que la utilización de la información generada está muy lejos de ser la óptima, conclusión a la que también arribaron las metaevaluaciones realizadas en otros contextos nacionales (véase capítulo 6).

#### Análisis FODA del sistema de evaluación

Este análisis se basa en la consideración de aspectos internos o endógenos (fortalezas y debilidades) y externos o exógenos (oportunidades y amenazas) de un sistema cualquiera.

Aplicado al sistema de evaluaciones externas revisadas, obtenemos los siguientes resultados principales:

- Fortalezas:
  - Prácticamente no existe en toda AL un esfuerzo de evaluación sistemático y obligatorio de los programas gubernamentales como el que se está llevando a cabo en México. Esto está generando un mercado, inexistente hace cuatro años atrás, que ha contribuido a desarrollar nuevas habilidades tanto del lado de la demanda (precisión en la elaboración de acuerdos, bases, términos de referencia, control de informes, de pagos, etcétera) como de la oferta de profesionales, que casi no contaban ni con una formación especializada en la materia ni con experiencia empírica.
  - Con la aprobación en 2002 de Ley Federal de Transparencia y Acceso a la Información Pública Gubernamental y la creación del IFAI se produjo un importante avance en materia de rendición de cuentas ya que, al menos es posible solicitar y obtener los informes de las evaluaciones realizadas de los programas que no se publican en Internet, así como sus respectivos costos.
  - También la LGDS, aprobada en 2003, constituye un importante avance para la continuidad y disposición de recursos para los programas sociales (artículo 20), el reconocimiento de la Contraloría Social como el mecanismo de participación de los beneficiarios (artículo 69 a 71) y, en materia de evaluación, la creación del CNE (artículo 81 a 85) que coordinará el sistema en su conjunto y en el que participarán investigadores académicos de probada calidad.
- Debilidades:
  - Mientras otros países que iniciaron procesos de evaluación, incluso más acotados que el mexicano, prepararon el terreno mediante cursos de capacitación, asesorías, elaboración de guías para realizar evaluaciones, etcétera,<sup>100</sup> y los pusieron en marcha en forma incremental,<sup>101</sup> México se lanzó en 2001 a la evaluación anual de todos los programas sujetos a reglas de operación en relación con su contexto, procesos, resultados e impactos,<sup>102</sup> lo que ha provocado el surgimiento y per-

<sup>100</sup> Sería el caso de Reino Unido y Francia, y también de los cuatro países latinoamericanos revisados (Colombia, Costa Rica, Uruguay y Chile).

<sup>101</sup> Salvo Uruguay, en el conjunto de los países antes nombrados.

<sup>102</sup> No se conoce ningún otro país que pretenda hacer algo similar; en particular, especialistas como Thoenig y Nioche señalaron que ningún país de la Unión Europea lo ha intentado.

manencia de múltiples problemas (que se analizarán en el siguiente apartado), algunos de los cuales requerirán del mediano plazo para su posible solución. La dificultad no radica sólo en la magnitud del trabajo propuesto sino en su factibilidad (disponibilidad de recursos de todo tipo que permitan cumplirlo), y fundamentalmente, si tiene sentido realizarlo en términos de la calidad de los resultados obtenidos y los procesos de aprendizaje iniciados.

- La evaluación de resultados revisada incurre en problemas metodológicos importantes en su aplicación, especialmente en relación con la representatividad de las muestras utilizadas y la emisión de juicios valorativos fundamentados a la gestión del programa evaluado. Tampoco se ha avanzado en el diseño de indicadores que sintetizen la información de la evaluación en correspondencia con los aspectos demandados en el Acuerdo.
  - La medición de impacto es el verdadero corazón de la evaluación de programas sociales, ya que entre la identificación de sus acciones y resultados y su balance de efectos pueden actuar diversas variables que provocan brechas que alejan los impactos reales de los esperados según su lógica interna. Sin embargo, en la práctica de las evaluaciones externas, no se realiza generalmente la evaluación de impactos o los métodos experimentales empleados dejan muchas dudas. Por otro lado, los trabajos de Sedesol miden la evolución de la pobreza, pero no se sabe si ésta obedece a los programas de la estrategia Contigo o a otras razones como el envío de remesas por parte de los trabajadores mexicanos en EUA (BM, 2004a).
  - La evaluación externa podría estarse transformando en un procedimiento más, preocupado de no perjudicar a nadie, un mero ritual como Lacasse señaló en el caso canadiense, o la institucionalización de un conjunto de prácticas estereotipadas, perdiendo su creatividad y adecuación a cada programa concreto.
- Amenazas:
    - Existe el riesgo de que, en un afán por estar al corriente en materia de modernización administrativa, se ponga una ambición exagerada sobre las posibilidades reales de la evaluación en nuestro contexto y se la encauce con un enfoque poco apropiado.
    - Sin un fuerte liderazgo, compromiso y participación en los sistemas de monitoreo y evaluación, el Banco Mundial (2004a) advierte sobre los riesgos de que no se logre su perfeccionamiento.

- Si la evaluación no demuestra rápidamente su utilidad podría truncarse el proceso o bien dedicársele recursos presupuestales más limitados.
- La falta de participación social en la evaluación puede interpretarse como ausencia de espíritu democrático y restar credibilidad y legitimidad a sus resultados.
- Oportunidades:
  - México cuenta con la gran oportunidad de mejorar el proceso iniciado, resolviendo sus problemas pendientes, integrando sus resultados a la gestión y difundiendo ampliamente, de manera de contribuir a la racionalización del gasto al mismo tiempo que al desarrollo democrático.

En el siguiente apartado se analizarán estos problemas para ordenarlos y jerarquizarlos, de forma de buscar soluciones en forma efectiva.

### Árbol de variables y problemas identificados

A partir de la larga lista de problemas identificados durante el trabajo realizado en los capítulos previos, se utiliza un diagrama de árbol para ordenar las variables en las que se manifiestan problemas y resaltar sus interrelaciones teóricas<sup>103</sup> (véase diagrama 3).

De esta forma aparece en el centro del diagrama la variable relacionada con el principal problema a resolver: *calidad de las evaluaciones*. Aunque en el capítulo 11 se identificaron evaluaciones de buena calidad, pudo verse que en 16 de los 31 programas evaluados no se cumplían los requisitos mínimos seleccionados para que éstas fueran consideradas aceptables. Además, en los tres casos en que se analizó la evolución de las evaluaciones se detectó una tendencia hacia el deterioro de su calidad en el tiempo.

Dicha calidad aparece afectada principalmente por otras seis variables que también presentan problemas:

#### Exigencias del Congreso

Nos referimos especialmente a la exigencia, incluida en el PEF, de evaluar todos los programas sujetos a reglas de operación, todos los años y en todos sus aspectos, lo que constituye una solicitud muy ambiciosa

<sup>103</sup>Cada nivel es afectado por variables del anterior (graficadas abajo) y afecta a las presentadas en el siguiente (graficadas arriba).

en cualquier contexto nacional y especialmente cuando estos procesos se están iniciando y no se cuenta con los recursos imprescindibles para implementarla.

#### Acuerdo SHCP-SFP

Este acuerdo establece los requisitos mínimos que deben cumplir los equipos de evaluación y los informes que elaboren, por lo que sirve de fundamento a la redacción de las bases y términos de referencia. El mismo solicita algunos aspectos prácticamente imposibles de cumplir con los sistemas de información, recursos y tiempos actualmente disponibles, como medir el impacto o la cobertura y su evolución por núcleo de población, además de presentar algunas imprecisiones.

#### Expectativas de Sedesol

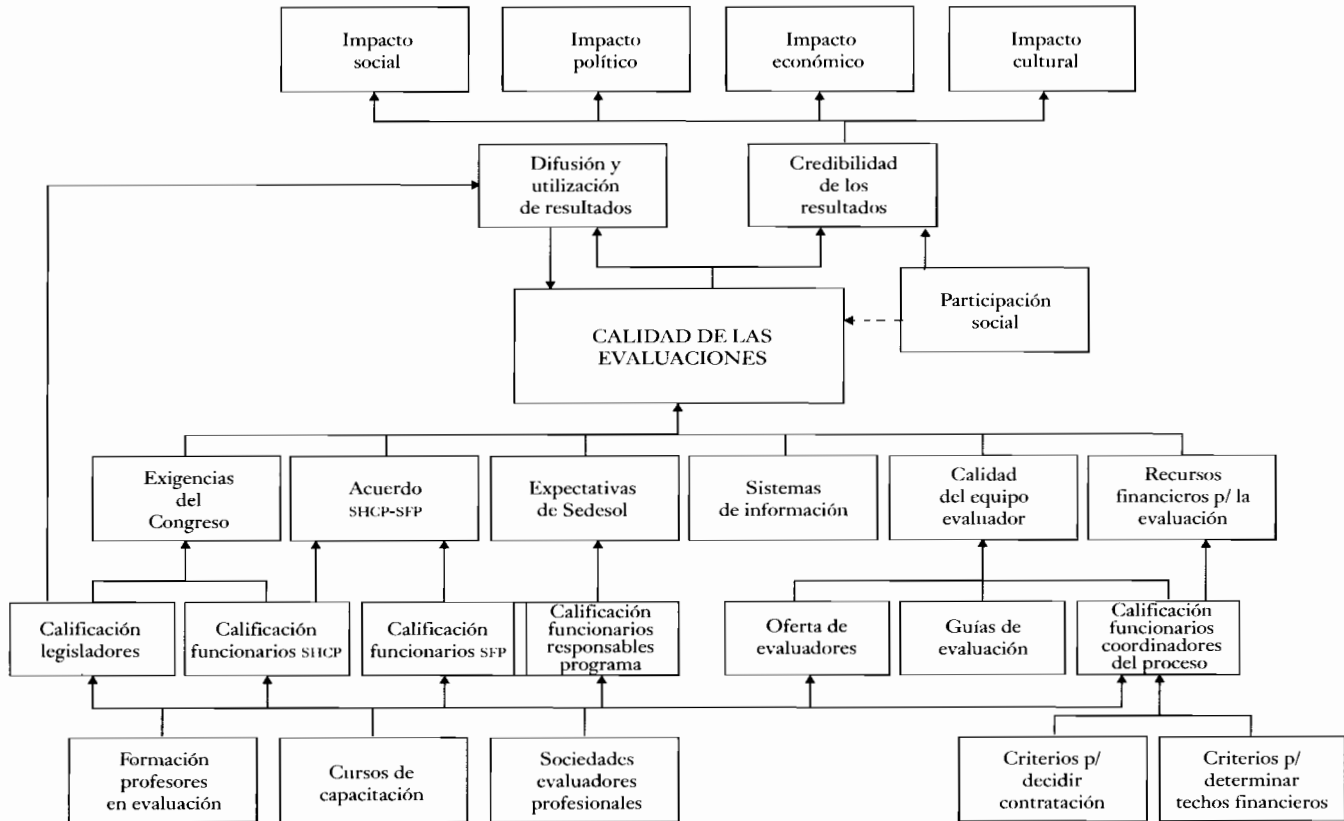
Como ya se mencionó, las expectativas de Sedesol se enfocan a la realización de estudios experimentales o cuasiexperimentales con mediciones econométricas en todos los programas. Pero, esto tiene dos problemas: sus costos y la poca credibilidad de sus resultados. En consecuencia, se considera necesario reflexionar sobre el grado de conveniencia de continuar promoviendo esta vía, particularmente discutida en el caso de los programas sociales. Actualmente, aunque los estudios experimentales no se logren por falta de recursos, aparecen como el tipo de ideal a alcanzar, en lugar de intentar buenos trabajos por otros medios. Además no está claro que constituyan la mejor opción a juzgar por los resultados de experiencias similares desarrolladas en EUA y abandonadas definitivamente en Francia.

#### Sistemas de información institucional

Estos sistemas de las instancias operadoras deben brindar información suficiente, oportuna, consistente y confiable, que sirva de partida a la evaluación, para su posterior complementación y verificación a través del trabajo de campo; lo que no sucede en nuestra realidad. La mayoría de los programas analizados no cuenta con padrones de beneficiarios que sirvan de insumo a la determinación de la muestra a encuestar, ni tampoco con datos agregados de todo el país y su distribución por estados y municipios que permitan al evaluador su verificación por muestreo. Sólo así pueden medirse variables como la cobertura y controlarse los costos de las evaluaciones, cuyos precios llegan a diferir de los informados por los propios evaluadores.

DIAGRAMA 3

### RED DE VARIABLES IMPLICADAS EN LOS PROBLEMAS DE EVALUACIÓN





### Calidad del equipo evaluador

Determinada por su formación y composición interdisciplinaria y *ad hoc* para el programa a evaluar, e integrada por especialistas con experiencia en la materia. Estas exigencias sólo resultaron cubiertas en ocho de los 18 casos en que se contó con la información correspondiente.

### Recursos financieros destinados a la evaluación

Es importante que los recursos sean congruentes con la magnitud del programa y el trabajo que implica su evaluación; sin embargo, existen algunas evaluaciones cuyo costo no parece justificarse y muchas que no cuentan con suficientes recursos para el trabajo que se les demanda. En 2002 se observa que el costo más alto supera 192 veces al más bajo y que estos montos tampoco están correlacionados con los presupuestos de los programas evaluados, variando la proporción Costo de la Evaluación/Presupuesto Ejercido por el Programa entre 0.03 y 5.11 por ciento. Tampoco existe correlación entre el costo y la calidad de la evaluación en 15 casos, 10 de ellos con exceso de costo y otros cinco con calidad superior a éste. Además, existe una tendencia al crecimiento no justificado de la más cara de ellas, que coincide con su asignación por vía directa.

A su vez, algunas variables del nivel que se acaba de analizar son afectadas por otras vinculadas a la calificación de diferentes grupos de recursos humanos en materia de evaluación, la existencia de guías de apoyo para su realización y la oferta numérica de evaluadores:

### Calificación de los legisladores

Esta variable determina en buena medida las ambiciosas exigencias de la Cámara de Diputados, plasmadas en los PEF, y contribuye a explicar la razón de que los congresistas no tengan capacidad de análisis de los informes (salvo la ASF, que hace sus propias auditorías técnicas de desempeño) y no estén utilizando sus resultados a pesar de ser los responsables formales de la entrada en vigencia de la obligación de evaluar. Lo anterior a pesar de que el PEF 2004 establece en su artículo 55 que las evaluaciones serán responsabilidad de la Cámara de Diputados, respecto de lo que no se perciben cambios.

### Calificación de los funcionarios de la SHCP

Dicha calificación es alta en materia de evaluación financiera pero no está especializada en la medición de impactos de programas sociales, lo que probablemente incidió para que incluyera en el proyecto de presupuesto para 2001 y siguientes, una propuesta demasiado ambiciosa, que la Cámara de

Diputados ha transformado en exigencia del PEF. Lo mismo podría haber sucedido con el Acuerdo SHCP-SFP.

#### Calificación de los funcionarios de la SFP

Su calificación ha estado históricamente referida a la realización de auditorías financieras y al control de la legalidad del gasto y la inversión pública, pero tampoco está especializada en la medición de impactos sociales, lo que dificulta su trabajo tanto para el control de la calidad de las evaluaciones como para el seguimiento del Acuerdo.

#### Calificación de funcionarios responsables del proceso

Puesto que los funcionarios de la Dirección General de Evaluación de la Sedesol son los responsables de marcar el rumbo de los procesos de evaluación, se considera ésta una variable de trascendental importancia. Si bien ellos cuentan con un alto grado de calificación en la materia, sorprende que insistan en defender la representatividad de las pequeñas muestras determinadas sin rigor estadístico y la contundencia de los resultados que de su aplicación se desprenden.

Como se planteó en la primera parte de este libro, la evaluación puede asumir dos corrientes diferentes: enfatizar la medición de impactos por medio de métodos experimentales de fuerte contenido cuantitativo y trasfondo positivista o dar prioridad a la formación de capital humano por vía del aprendizaje basado en la participación, con un diseño constructivista, que utiliza información cuantitativa menos rigurosa y la combina con estudios cualitativos que profundizan en la comprensión de los problemas.

El problema que se percibe es que, a pesar de que los funcionarios de Sedesol dicen valorar el trabajo cualitativo, su énfasis es netamente cuantitativo y su aspiración es llegar a aplicar estudios experimentales en todos los programas en que los recursos financieros y humanos calificados se los permitan, en consistencia con su formación predominante de carácter económico-matemático.

Sin embargo, la aplicación de dichos métodos en el caso de Progres-Oportunidades no ha permitido aún, a pesar de los cuantiosos recursos invertidos en su evaluación, contestar algunas de las preguntas que se planteaba Skoufias o que mencionan algunos de los entrevistados en este trabajo, como: ¿cuál es el impacto sinérgico logrado por la oferta conjunta de servicios de educación, salud y alimentación, frente a los programas que los ofrecen por separado?, ¿se ha producido un incremento de la calidad de la educación y del aprendizaje?, ¿cuál es el monto mínimo a transferir para lograr la perma-

nencia de los niños en la escuela y el acceso regular a los centros de salud? o las mejoras reportadas por el gobierno en materia de pobreza extrema, ¿realmente fueron provocadas por Oportunidades o la estrategia Contigo en su conjunto?

Los propios evaluadores del programa señalan que los resultados alcanzados son indicativos, que ameritan estudios en mayor profundidad, análisis más detallados, mediciones más precisas, que los datos “sugieren”, las conclusiones “apuntan a” o simplemente que “se espera que” se traduzcan en mejores oportunidades futuras de trabajo y condiciones de vida. Este tipo de conclusiones, tan lejanas de la relación causa-efecto de las ciencias “duras” a las que se tomó prestado el modelo, deberían hacer reflexionar sobre la conveniencia de su utilización y la justificación de su costo.

#### Oferta de evaluadores

A mayor cantidad de evaluadores en el mercado, mayor posibilidad de seleccionar los mejores equipos; sin embargo, existe insuficiente número de evaluadores con formación y capacitación adecuada para hacer trabajos de calidad.

#### Guías de evaluación

No se ha realizado ninguna guía de evaluación para apoyar la calidad del trabajo de los evaluadores, a diferencia de las varias que circulan en países avanzados en la materia como Reino Unido y Francia, y también en algunos latinoamericanos. En menor medida, éstas también podrían ser de utilidad para los legisladores y funcionarios relacionados con los procesos de evaluación.

#### Calificación de funcionarios coordinadores del proceso

Nos referimos a los funcionarios que determinan los techos financieros dedicados al pago de evaluaciones y los que componen las oficinas de las instancias operadoras dedicadas a la relación con los equipos de evaluadores, en materia de llamados a licitación, decisión de los ganadores del concurso en comisiones *ad hoc*, control de tiempo y productos, autorización de pagos, entrega de informes, etcétera. En opinión de algunos entrevistados y en función de su formación, muchos de ellos no tendrían la calificación necesaria para cumplir con éxito sus actividades, lo que provocaría irregularidades en el proceso. Además, se detecta un conflicto de intereses que se traduce en presiones más o menos sutiles sobre el evaluador al ser la misma instancia la que opera el programa, maneja el proceso de evaluación y, fundamentalmente, paga el servicio con sus propios recursos.

Nuevamente este nivel es explicado por otro conjunto de variables:

#### Formación profesional en evaluación

Esta formación depende de la cantidad y calidad de la oferta educativa realizada por las universidades. En el caso mexicano, los coordinadores de los equipos evaluadores cuentan con doctorados y maestrías en diversas áreas del conocimiento (economía, administración pública, sociología, antropología, etcétera), pero ninguno ha realizado un posgrado específico en evaluación, mismo que tampoco se imparte en el país.

#### Capacitación en evaluación

El tipo de cursos y seminarios que Sedesol desarrolla se enfoca especialmente al aprendizaje de modelos experimentales y mediciones econométricas. Esta elección resulta totalmente opuesta a la que podría orientar la desilusión que estos mismos métodos han generado en su país de origen. También resulta muy diferente de las preocupaciones que se incluyen en un taller de desarrollo profesional organizado recientemente por la European Evaluation Society,<sup>104</sup> donde los debates sustantivos de la práctica se priorizan sobre el aprendizaje de métodos y, más aún, de la estadística aplicada, a la que ni siquiera se incluyen referencias. Además, nuestras universidades ofertan muy pocos cursos de capacitación en evaluación.

#### Sociedades de evaluadores profesionales

Las asociaciones profesionales constituyen espacios en que se debaten temas especializados, se ofrecen cursos, asesorías o consultorías. Las mismas existen en Francia y la mayoría de los países europeos, EUA y Canadá, pero aún no se desarrollan en nuestro contexto.

<sup>104</sup> Dicho programa, desarrollado en Dinamarca en octubre de 2003, incluyó:

- relevancia de la teoría en la práctica evaluativa: teoría del programa y teoría de su implementación; teorías de los tomadores de decisión, implementadores, beneficiarios y otros grupos de interés; teoría y solicitudes de evaluación; teorías de diseño evaluativo y selección de métodos;
- instrumentos de política y evaluación: tipologías de instrumentos de políticas; metas y teorías de instrumentos de políticas; instrumentos de políticas y teorías de programa; implicaciones para los diseños evaluativos;
- evaluación y políticas basadas en evidencia: mecanismos, contexto, resultados; síntesis realista de lo que puede ser comparado;
- el impacto de la evaluación en la sociedad y la actividad humana; reflexión y aprendizaje: el proceso de evaluación; evaluación como educación;
- ¿qué podemos aprender?: aspectos importantes en la práctica; formas de manejarlos; formas de ayudarnos a trabajar mejor.

### Criterios legales para decidir la forma de contratación

En el contexto mexicano deberían definirse criterios que restringieran los casos en que es posible contratar los servicios de instituciones públicas para realizar evaluaciones sin licitación previa, puesto que la Ley de Adquisiciones, Arrendamientos y Servicios del Sector Público sólo es aplicable para regular las contrataciones con el sector privado. Por otra parte, esta Ley enfatiza los procedimientos de compra de bienes y ciertos servicios, pero no se adecua a la contratación de trabajos de consultoría y asesoría; en particular, exige otorgar el contrato a quien, cumpliendo los requisitos, ofrezca el precio más barato, lo que no permite valorar diferencias de calidad en el trabajo a realizar. Los problemas anteriores dificultan el trabajo de los funcionarios que deben coordinar el proceso.

### Criterios para determinar techos financieros para la evaluación

Estos criterios constituyen un insumo fundamental para guiar las tareas de los funcionarios responsables de la determinación del techo financiero establecido por la instancia operadora y pueden definirse en función de la fijación de precios mínimos y máximos, como porcentaje del presupuesto, etcétera. Hasta el momento no se conoce ninguna normatividad que los regule.<sup>105</sup>

A continuación analizaremos los efectos producidos por la variable Calidad de las Evaluaciones:

### Difusión y utilización de resultados en la gestión

A pesar de su debilidad, se difunden ampliamente los buenos resultados alcanzados, sin mencionar los impactos negativos, como los que destaca la evaluación cualitativa de Oportunidades, muy coincidentes con los hallados realizados por las ONG. Una difusión más justa debería poner a disposición del público los informes completos, con su balance de resultados para ser utilizados en la toma de decisiones.

Los resultados de las evaluaciones no se usan para tomar decisiones en el Congreso y se utilizaban poco en 2002 en la Sedesol (para mejorar términos de referencia, principalmente) y las instancias operadoras (para mejorar algún aspecto de su gestión), por lo que se ha corrido el riesgo de transformar el proceso en un rito inútil.<sup>106</sup>

<sup>105</sup> El director de metodología de la Sedesol mencionó que el precio debe ser inferior al 7 por ciento de los gastos indirectos del programa, pero se desconoce en qué norma estaría establecido dicho límite.

<sup>106</sup> Esto estaría cambiando desde 2003, en que han dejado de utilizar como indicador interno el porcentaje de programas evaluados (100 por ciento) y lo han sustituido por el porcentaje de recomendaciones externas atendidas por el programa, según informó el director general de evaluación de Sedesol.

Esta variable, como ya se mencionó, también es consecuencia de una analizada en el segundo nivel (calificación de los legisladores) y a su vez, impacta a la Calidad de las Evaluaciones al desmotivar a los evaluadores en la realización de buenos trabajos ya que estos saben que tienen alta probabilidad de no ser conocidos ni tomados en cuenta.

#### Credibilidad de los resultados

Ésta depende de la metodología empleada en la evaluación, pero también de rasgos de honestidad y honradez probada de los evaluadores. Como ya se ha mencionado, dicha metodología deja muchas dudas sobre la confiabilidad de los datos y sus métodos de análisis, mientras se desconoce el comportamiento ético de la mayoría de los evaluadores. Además, esta variable es afectada por otra del conjunto:

#### Participación social en la gestión y evaluación

Ésta puede contribuir a lograr resultados más efectivos en ciertas circunstancias, pero también sirve para generar legitimidad, regulación por cooperación, gobernanza y aprendizaje a través de sus procesos. A pesar de que los planes y programas gubernamentales la prevén, los evaluadores coinciden en que los mecanismos planeados (Contraloría Social, consejos Microrregionales, etcétera) en algunos programas no funcionan. Parece haber muchas dudas sobre su efectividad y, tal vez, temor al “empoderamiento” de los beneficiarios y las OSC entre quienes deben poner en marcha los mecanismos planeados. Sin embargo, también existen muchos desarrollos teóricos y experiencias exitosas que consideran que su desempeño ha sido incrementado por la participación social, por lo que la relación de esta variable con la calidad de las evaluaciones se deja pendiente hasta que se realice un mayor número de investigaciones rigurosas.

Finalmente, las dos variables afectadas por la calidad de las evaluaciones provocan a su vez impactos en el contexto de orden social, político, económico y cultural:

#### Impacto social

En la medida en que las evaluaciones son muy poco utilizadas para mejorar los programas sociales, éstas no se traducen en un mayor impacto perceptible hacia la solución del problema que les dio origen. Tampoco generan un aprendizaje entre los beneficiarios y las OSC por falta de participación. Muchos funcionarios y beneficiarios encuestados expresan su desilusión porque los datos

y opiniones que manifiestan durante el trabajo de campo de la evaluación no se traducen en modificaciones en el diseño y la gestión del programa.

### Impacto político

Si los informes de evaluación se difunden en forma muy restringida, estos no pueden provocar impactos importantes en términos de la legitimidad del gobierno y sus acciones. En cambio, se realiza mucha propaganda política en los medios que, como ya se mencionó, destaca algunos aspectos positivos de su gestión. Pero debido a que los mismos gozan de poca credibilidad, habría que hacer un estudio específico para conocer con precisión cuáles son sus efectos en la ciudadanía.

### Impacto económico

En 2002, el costo de las evaluaciones insumió 52'546,000 pesos que corresponde a una pequeña fracción de los recursos disponibles en los presupuestos de los programas sociales (0.26 por ciento).<sup>107</sup> Por otra parte, podría estarse produciendo un impacto económico en la distribución de los recursos presupuestales entre los distintos programas a través de la elaboración del proyecto de presupuesto por la SHCP. Esto no sucede a nivel de la Cámara de Diputados, puesto que ésta no utiliza la información de las evaluaciones.

### Impacto cultural

Poco a poco los valores, conocimientos, tradiciones y creencias que provocaron décadas de rechazo al tema de la evaluación en todos sus tipos y enfoques, están siendo modificados por los procesos en curso, especialmente en el marco de los programas sociales del gobierno y de las OSC. Sin embargo, aún queda mucho por avanzar en ese camino, y sus posibilidades y ritmos dependerán de la difusión, la credibilidad y la demostración de la utilidad de tomar en cuenta sus resultados en la gestión de los programas.

Hasta aquí el diagrama de árbol nos ha ayudado a organizar las variables en que se presentan problemas y establecer sus interrelaciones; a continuación nos permitirá identificar dentro del sistema amplio en análisis las variables-problema que aparecen como las raíces del problema central, aquellas que, en buena medida, aparecen determinadas en forma exógena al sistema mismo porque dependen de decisiones asumidas por instancias o autoridades relativamente ajenas a las involucradas directamente en la práctica de la evaluación.

<sup>107</sup> En el caso de los programas de BM esta relación varía entre el 2 y 3 por ciento .

No se pretende que las variables afectadas por las causas iniciales sean determinadas exclusivamente por éstas, pero sí que mejorar el comportamiento de las primeras provocaría una mejora sustantiva en las últimas, hasta alcanzar a la vinculada al objetivo central del árbol: la calidad de las evaluaciones.

Algunas medidas pueden comenzar a incidir en el corto plazo si se cuenta con la disposición de la instancia responsable, pero otras como la capacitación requieren como mínimo un año, y la formación universitaria alrededor de tres entre la planeación y el egreso de una primera generación de posgraduados.

En conclusión, las variables identificadas en el árbol como causas iniciales y las instancias responsables de adoptar decisiones para resolver sus problemas son:

- Formación profesional en evaluación: universidades.
- Cursos de capacitación en evaluación: universidades, secretarías y Congreso.
- Sociedades de evaluadores profesionales: evaluadores.
- Criterios legales para decidir la forma de contratación: Congreso.
- Criterios para determinar techos financieros para la evaluación: SHCP, Congreso.
- Sistemas de información institucional: organizaciones operadoras de programas, Sedesol.
- Participación social en la gestión y evaluación: organizaciones operadoras de programas, Sedesol, evaluadores.

En consecuencia, las responsabilidades recaen principalmente en las universidades y en las áreas de capacitación de la Sedesol y el Congreso, seguidas de las áreas de gestión y coordinación de la evaluación en las instancias operadoras de programas, la SHCP y los propios evaluadores.

El primero y el sexto de los problemas enunciados se refieren a dos de los tres requisitos fundamentales manifestados por Abate (2000) para hacer posible la puesta en práctica de los nuevos modelos de gestión: amplia base profesional con conocimientos y capacidades desarrolladas que garanticen las iniciativas gerenciales, y posibilidad de los sistemas de información de medir costos, productos, calidades e impactos de las acciones realizadas.

## Propuestas

A continuación se presentan propuestas para resolver los problemas básicos localizados en el árbol de análisis previo y algunas sugerencias inmediatas



que puedan mejorar en forma más directa la calidad de las evaluaciones. Se plantea para las primeras:

- **Formación profesional en evaluación.** Esta calificación se debe realizar para evitar improvisaciones; alude a la formación previa al desempeño de la función y requiere urgentemente el diseño de posgrados universitarios en evaluación de programas, con especialización en los de índole social, conocimiento de los distintos enfoques y métodos disponibles y formación fuertemente interdisciplinaria.<sup>108</sup> Se sugiere un esfuerzo combinado de varias instituciones de educación superior de prestigio reconocido.
- **Cursos de capacitación en evaluación.** Nos referimos a la capacitación continua durante el trabajo, mediante el desarrollo articulado de múltiples seminarios, congresos, cursos, etcétera, que mejoren la calificación de los funcionarios actuales y desarrollen un currículum similar al esbozado en el punto anterior. Se recomienda un esfuerzo conjunto de las universidades con el Congreso y la Sedesol, recuperando la experiencia del gobierno británico en la materia.<sup>109</sup>
- **Sociedades de evaluadores profesionales.** Estas sociedades han contribuido en forma importante tanto a la calificación de funcionarios como a dar seriedad a los procesos, encargándose incluso en algunos países de certificar a los evaluadores. También han permitido llevar el conocimiento y los debates a las diferentes regiones a través de la creación de filiales como sucede en el caso de la SFE y reunir a las diferentes naciones de un continente como en el caso de European Evaluation Society. Se recomienda promover su conformación entre los evaluadores más reconocidos.
- **Criterios legales para decidir la forma de contratación.** Se sugiere a los legisladores promover una ley específica para la contratación de servicios de consultoría y asesoría que prevea con claridad las situaciones que ahora resultan inadecuadas y probablemente ineficientes.

<sup>108</sup> Para Stufflebeam y Shinkfield (1987: 25) los evaluadores deben contar con conocimientos sobre "... las entrevistas, los informes preliminares, el análisis de contenidos, la observación, el análisis político, el análisis económico, los informes técnicos, el estudio de caso, la evaluación sin metas, la escucha de opiniones contrapuestas, las asociaciones de defensa, las listas de control, la elaboración de pruebas, el análisis estadístico, el diseño de la investigación, el análisis sistemático, la teorización y la administración del proyecto".

<sup>109</sup> Díaz y Rosales (2003) también reconocen la ausencia de programas de formación y capacitación, por lo que recomiendan la creación de una escuela permanente, que concientice a los participantes en el tema, enseñe a trabajar en equipos multi e interdisciplinarios, e imparta teoría y métodos de evaluación, pero también aporte un mayor conocimiento, de carácter sistémico, no especializado, del campo específico de los programas evaluados.

- Criterios para determinar techos financieros para la evaluación. Estos podrían ser propuestos por la SHCP en el próximo proyecto de PEF para su discusión y, en su caso, aprobación por parte de la Cámara de Diputados.
- Sistemas de información institucional. Urge el diseño o mejoramiento de los actuales sistemas de información de manera de contar con padrones de beneficiarios de todos los programas sujetos a reglas de operación, así como con información suficiente y confiable de los presupuestos ejercidos, las acciones y los servicios desarrollados por el programa en las diferentes regiones en que el mismo se aplica.
- Participación social en la gestión y evaluación. Mucho se ha escrito heurísticamente sobre las bondades de la participación social en la efectividad y legitimidad de los programas y de su evaluación y también muchos especialistas han argumentando su inutilidad. En materia de evaluación, ella permite comprender los problemas tal cual los viven los grupos sociales afectados y evaluar el diseño de programas en función de ello (como complemento de la visión burocrática), así como facilita la introducción de nuevas estrategias por actores motivados. Pero no sólo puede mejorar la calidad técnica de los resultados de la evaluación (al incluir las miradas y valoraciones de diferentes actores), sino contribuir sustantivamente a su credibilidad y legitimidad.

Es obvio que hay experiencias de participación exitosas y fallidas por lo que no es posible prescribirla en todos los casos, pero los planes y programas gubernamentales han optado por incluirla, por lo que se propone poner en marcha mecanismos efectivos para la integración de los organismos participativos previstos por la Ley y los programas, así como solicitar algunas evaluaciones piloto de tipo participativo, mediante una estrategia incremental (pasando de una etapa a la siguiente sólo después de haber tenido éxito en la anterior):

1. en lo inmediato, tomar en cuenta por parte de las autoridades las opiniones de los ciudadanos en torno a la orientación del programa, su organización, gestión y evaluación e informarles de los resultados;
2. a corto plazo, tomar decisiones y evaluar resultados en forma conjunta entre autoridades y ciudadanos; y
3. a mediano o largo plazo, permitir a los últimos el liderazgo y la administración de los recursos del programa.

La única forma de poner fin a un debate que no cuenta con información rigurosa es someter a prueba la hipótesis mediante el desarrollo de una investigación social que permita analizar si hay o no diferencias significativas en la gestión y evaluación participativas y en qué condiciones se presentan.

A mediano plazo, y una vez mejoradas las calificaciones actuales, se podría considerar la posibilidad de que el propio Congreso y sus comisiones se encarguen de solicitar evaluaciones externas,<sup>110</sup> en su calidad de contrapeso del Poder Ejecutivo (como la GAO en EUA o la NAO en Reino Unido), aunque existe el riesgo de que éstas se politicen.

Incluimos como propuestas dirigidas a incidir de inmediato en los trabajos de evaluación:

- Continuar con el proceso de integración de programas con objetivos similares iniciado en Sedesol en 2003,<sup>111</sup> y promover su aplicación a nivel intersecretarial, evitando duplicidades y reduciendo la proliferación de pequeños programas que dificultan y hacen costosa su evaluación por separado.
- Impulsar la celebración de un nuevo Acuerdo SHCP-SEF para clarificar conceptos básicos (evaluación, sus objetivos y tipos, cobertura, impacto, métodos, criterios, por ejemplo), ampliar requisitos (evaluar la pertinencia y el diseño del programa,<sup>112</sup> y analizar la participación social en su gestión) y simplificar procedimientos poco realistas (medición de impactos anuales o diseño de escenarios prospectivos). Se sugiere solicitar también la distribución de resultados por microrregiones de manera de realizar una especie de *benchmarking* para determinar los casos exitosos, sus costos, las diferencias con los no exitosos y sus causas, para tomar medidas que permitan una mejora generalizada. Además y fundamentalmente, agregar la obligación del evaluador de incluir en su informe final un juicio explícito de valoración fundamentada de la gestión, antes de pasar a las recomendaciones.
- Modificar el contenido de los decretos recientemente expedidos correspondientes a la reglamentación de la LGDS y la regulación del CNE de manera que se respete plenamente lo establecido por la mencionada

<sup>110</sup>El PEF 2004 establece en su artículo 55 que las evaluaciones serán responsabilidad de la Cámara de Diputados, sin embargo, no se perciben cambios en su organización.

<sup>111</sup>Caso de Desarrollo Económico Productivo de los Pueblos Indígenas y Desarrollo de los Pueblos y Comunidades Indígenas, ambos vigentes en 2002.

<sup>112</sup>Algunos programas comenzaron a hacerlo en 2003.

ley, en aras de la claridad y legitimidad de los procesos de gestión y evaluación de los programas de desarrollo social.<sup>113</sup>

- Impulsar las modificaciones a la LGDS que sean necesarias para: 1. crear un organismo autónomo del gobierno y del Congreso<sup>114</sup> que se encargue de hacer las evaluaciones o de coordinar la participación de los académicos desde la licitación hasta la entrega, difusión y publicación de sus resultados, evitando así los conflictos de intereses y posiciones partidistas;<sup>115</sup> y 2. realizar anualmente evaluaciones de insumos, procesos y resultados de todos o una muestra de los programas<sup>116</sup> y limitar la evaluación de impacto de acuerdo con alguna o algunas de la siguientes sugerencias:
  - Medir sólo el impacto del conjunto de programas sociales a nivel de la estrategia Contigo. De esta forma, habría que estudiar únicamente la evolución de los indicadores de pobreza y descartar la posible incidencia de variables contextuales, ajenas a la estrategia, sobre dichos resultados. Se evitaría así la casi imposible tarea de deslindar los impactos causados por cada programa, cuyos objetivos y poblaciones suelen trasladarse, y se contaría con posibilidades más realistas de ofrecer verdaderas mediciones de impacto.
  - Desagregar los indicadores anteriores con criterios territoriales, de manera de analizar el impacto que el conjunto de programas sociales ha tenido en cada una de las microrregiones, para hacer un *benchmarking* entre ellas y así contribuir a la evaluación de la estrategia del mismo nombre.
  - Evaluar selectivamente los impactos de algunos programas, de manera que transcurran por lo menos tres años entre una evaluación y la siguiente.<sup>117</sup> Esto permitiría que los efectos propios del programa

<sup>113</sup> Consultar artículos publicados por Julio Boltvinik en el periódico *La Jornada* de fechas 10, 17 y 24 de febrero y 3 y 10 de marzo de 2006, entre otros.

<sup>114</sup> La iniciativa Burgos de LGDS preveía la creación de un órgano de ese tipo. Instituto Nacional de Evaluación de la Política Social, de carácter autónomo, similar al Instituto Federal Electoral, al Banco de México y al Instituto Nacional de Evaluación de la Política Educativa, creado en agosto de 2002.

<sup>115</sup> Su creación constituye un requisito fundamental para el éxito de la gerencia pública, según palabras de Weissbluth (2004), consultor y miembro del Consejo de Alta Dirección Pública de Chile. En el contexto español también acaba de aprobarse la creación de un organismo responsable de las evaluaciones. En este caso, si bien no es de carácter totalmente autónomo, depende de un ministerio transversal, el Ministerio de las Administraciones Autónomas, con lo cual se logra al menos que no esté directamente vinculado con ningún ministerio responsable del desarrollo de programas.

<sup>116</sup> La mayor parte de las evaluaciones francesas, por ejemplo, solicitan cubrir sólo cinco puntos: pertinencia del programa, coherencia del diseño, eficacia de las acciones realizadas, eficiencia en el empleo de recursos para el logro de objetivos, y conclusiones y recomendaciones.

<sup>117</sup> En su documento sobre la pobreza en México, el Banco Mundial (2004b: 238 y 247) expresa también esta idea de que la medición de impacto se aplique a los principales programas de desarrollo social, aquellos que tengan alta cobertura nacional y absorban recursos sustanciales, a proyectos innovadores, reproducibles y con

tengan mayor probabilidad de ser medidos, a pesar de sus interdependencias, y concentrar los recursos anuales disponibles (financieros y humanos) en una tercera parte de las actuales evaluaciones, de manera de hacerlas más realistas y rigurosas, recuperando la experiencia canadiense de finales de los ochenta en que las mismas se aplicaban cada cinco años.

- Simplificar la solicitud de evaluación de programas pequeños (con presupuestos inferiores a los 100 millones de pesos) que no ameritan el mismo rigor que el resto.
- Fortalecer las relaciones entre las instancias formalmente intervinientes en el proceso, comenzando por que las Comisiones de Desarrollo Social, Presupuesto y Cuenta Pública, y Participación Ciudadana del Congreso compartan y discutan al interior del mismo y con los funcionarios y evaluadores las auditorías de desempeño desarrolladas por la ASF y los informes de las evaluaciones externas. Retomar también la propuesta del BM (2004a) para conformar paneles de asesoría técnica con la participación de expertos en políticas, congresistas, beneficiarios y operadores de programas.
- Limitar anualmente a tres el número de evaluaciones en que pueden participar los mismos coordinadores y los miembros de sus equipos para garantizar la calidad de sus trabajos y ampliar la competencia en su mercado.
- Exigir que los informes de evaluación contengan un anexo en que se especifique y fundamente la forma en que fue determinada la o las muestras a encuestar, así como los parámetros utilizados en su cálculo y las limitaciones de la misma, de manera que el lector pueda valorar la calidad de la información que se le presenta.
- Aplicar una mezcla de enfoques cuantitativos y cualitativos integrados que permitan aprovechar al máximo las ventajas de cada uno de ellos: mayor alcance en extensión y en profundidad.
- Preparar resúmenes ejecutivos de la evaluación con contenidos diferentes según sus destinatarios: Sedesol, instancias operadoras de programas, SHCP, SFP, Congreso, ciudadanía (ejemplo: los legisladores necesitan información sobre los logros generales del programa para decidir sobre su continuidad y asignación de recursos, los operadores de programas requieren información más detallada

---

intervenciones bien definidas. La OCDE (1995) también recomienda llevar a cabo evaluaciones selectivas a los programas, de manera que los mismos resulten evaluados aproximadamente cada cinco años. Esto especialmente en los países en vía de desarrollo en que el número de buenas evaluaciones que es realista esperar resulta limitado.

sobre su eficiencia y calidad, la Sedesol requiere también información general sobre la eficacia y el impacto logrados de manera de poder realizar correcciones o un aprendizaje útil para futuros diseños, etcétera). Lo anterior permitiría facilitar su lectura debido a su menor extensión y concreción en los temas de interés de cada grupo de actores.

- Apoyar a los evaluadores externos por parte de las delegaciones estatales de las instancias operadoras del programa en la logística del trabajo de campo, pero dejarlos en libertad de acceder y seleccionar a los beneficiarios a encuestar, para que estos contesten con mayor libertad.
- Publicar los resultados de los informes de evaluación, no sólo en el *Diario Oficial* como lo prevé la LGDS aprobada en 2003, sino dar a conocer los resúmenes ejecutivos en las páginas web correspondientes a sus programas y difundir sus conclusiones en la prensa, como se hace en Costa Rica y Uruguay. Ello como requisito indispensable para que la ciudadanía esté informada y pueda participar democráticamente en su debate, lo que indirectamente promoverá una mejor calidad de los informes.
- Impulsar el desarrollo de trabajos de metaevaluación que ayuden a aprender de las experiencias desarrolladas y corregir el rumbo del sistema de evaluación en su conjunto.
- Comenzar a desarrollar una cultura de la evaluación que modifique valores, cambie actitudes y se traduzca en nuevas prácticas que sistematicen sus resultados y promuevan su utilización para fijar metas de impacto, de resultados, de actividades e insumos, en ese orden, y tomar decisiones para la conducción de la gestión, de manera de integrar la evaluación a la planeación y la operación, similar al sistema introducido en Canadá o la ABR que promueve el BM. Se requiere de un liderazgo que diseñe incentivos para impulsar el cambio cultural.

### Escenarios prospectivos

A finales de 2005 el Poder Ejecutivo federal expidió el reglamento de la LGDS y a inicios de 2006 se procedió a conformar el CNE.<sup>118</sup> Especialmente el último constituye un evento importante que puede modificar el futuro de la evaluación de programas.

<sup>117</sup> Según el segundo transitorio de la LGDS, la expedición del reglamento contaba sólo con 90 días naturales a partir de su entrada en vigencia, el 20 de enero de 2004; según el tercero transitorio de la misma, el CNE debía integrarse dentro de los 180 días naturales a contar desde la misma fecha. Este último quedó conformado el 9 de febrero de 2006.

Si el consejo mantuviera la situación que se ha venido presentando desde 2001, sin más cambios que los previstos por la LGDS (fechas de entrega, aprobación de un sistema de indicadores, publicación de informes, etcétera) tendríamos que esperar que el *escenario tendencial* de los próximos años estuviera caracterizado por:

- mantenimiento de muchas instancias vinculadas al proceso de evaluación, muy poco relacionadas, o mal relacionadas, entre sí,
- inercia en la expectativas y exigencias, en algunos casos demasiado ambiciosas, de Sedesol, SHCP y el Congreso;
- diseño de términos de referencia y procedimientos relativamente confusos;
- esfuerzos insuficientes de formación profesional y de capacitación de funcionarios y legisladores en el tema;
- poco incremento numérico de la oferta de evaluadores y de la calidad de sus equipos;
- tendencia de los anteriores a concentrar varios trabajos de evaluación, mismos que aún podrían disminuir su número si continúa el proceso de integración de programas;
- ausencia de guías o manuales de evaluación que apoyen su trabajo;
- sistemas de información insuficientes y poco confiables;
- continuación de la evaluación anual de todos los programas en todos sus aspectos, sin tiempo y recursos suficientes para hacerla con calidad y tendiendo a transformarla en un procedimiento más de control administrativo tradicional (ahora realizado por externos);
- levantamiento de información de campo con base en muestras de dudosa o desconocida representatividad;
- ausencia de evaluación de la pertinencia de los programas y de juicios valorativos fundamentados sobre su gestión;
- énfasis en la medición de variables económicas, en detrimento de las sociales, culturales y políticas;
- casi nula participación social en la gestión y evaluación de programas,
- ausencia de involucramiento del Congreso en los procesos de evaluación y aprobación de presupuestos no fundamentados en resultados;<sup>119</sup> y
- lentos cambios en la cultura organizacional en torno a la evaluación y las resistencias que dicho proceso desata.

<sup>119</sup>En 2004 la Sedesol habría empezado a tomar en cuenta los resultados de las evaluaciones para la definición de los presupuestos de los programas.

Sin pretender haber logrado un listado exhaustivo de características, todo lo anterior redundaría en una calidad de la evaluación que no mejora, incluso con el riesgo de empeorar en algunos casos.

Los informes continuarían difundiéndose en forma insuficiente, a pesar de la obligación introducida por la LGDS en el sentido de publicar sus resultados en el *Diario Oficial*, y utilizándose poco en la gestión de los programas, y por lo tanto, no se traducirían en impactos relevantes en lo económico, social, cultural y político.

Si todo esto sucediera, también se perdería la poca credibilidad actual del proceso y quedaría más clara su limitada justificación, con lo cual podría concretarse la amenaza de que el proceso se trunque, o más probablemente, que se le reduzcan los recursos.

En ese caso, continuaríamos manteniendo un esfuerzo que se transformaría en un rito sin sentido, pero seguiríamos sin conocer con rigor el impacto de las intervenciones sociales del gobierno.

Sin embargo, la Sedesol ha mostrado un gran interés en el avance de los procesos de evaluación. Por ejemplo, en la evaluación a la gestión 2003, suponemos que con la intención de exigir mayor interdisciplinariedad y calidad del equipo evaluador, los términos de referencia de algunos programas (Jóvenes por México, Iniciativa Ciudadana 3X1) empezaron a solicitar el cumplimiento de mayores requisitos por parte de los evaluadores, como:

- La dirección de la evaluación debía tener estudios de doctorado o equivalente y experiencia previa en evaluaciones; y
- el equipo tenía que contar con:
  - analistas (economistas y otros científicos sociales),
  - un experto en estadística,
  - un administrador del trabajo de campo y del equipo, y
  - procesadores de datos.

Estas medidas pueden ser discutibles ya que exigir el doctorado cierra mucho las posibilidades de la oferta, ya de por sí restringida, debido a muchos de los evaluadores actuales sólo cuentan con maestría. Además, si tener experiencia previa se refiere a la realización de evaluaciones externas anteriores, eso significaría que ningún evaluador nuevo puede ingresar al grupo, con lo cual la oferta se transformaría en una especie de oligopolio cerrado que no contribuiría a mejorar la calidad de las evaluaciones realizadas. De todas formas demuestra el interés de Sedesol y los operadores de los programas de buscar procesos de mejoramiento continuo.



Lo anterior nos hace suponer que es posible plantearse *escenarios alternativos* más optimistas.

En caso de que el CNE cuente con recursos suficientes y actúe con una visión de gestión estratégica, aprovechando las oportunidades y fortalezas del sistema de evaluación y corrigiendo sus debilidades o problemas para evitar las posibles amenazas del entorno, esperaríamos que introdujera todas o por lo menos algunas de las propuestas que se expusieron previamente, modificando la tendencia generalizada a la realización de estudios econométricos y provocando distintos grados de mejoramiento en la calidad de las evaluaciones, su difusión, utilización y credibilidad. En consecuencia, se esperaría un aumento sustantivo de su impacto en el contexto, en forma similar a la experiencia canadiense iniciada en el año 2000.

Para lograr revertir las tendencias del escenario tendencial, dicho consejo debería ejercer un fuerte liderazgo que iniciara un diálogo con las áreas de capacitación de las secretarías de Estado involucradas, las áreas que operan programas, las que coordinan sus evaluaciones y los propios evaluadores; pero puesto que, como ya se mencionó, no todos los problemas se pueden resolver al interior del gobierno mismo, sería también fundamental que incrementara sus vínculos con las universidades y el Congreso.

## Conclusiones

A modo de cierre se considera necesario hacer una síntesis de las conclusiones fundamentales que ha ido arrojando la metaevaluación realizada:

- No es conveniente evaluar individualmente cada programa para determinar su impacto en la medición de la pobreza porque los métodos experimentales dejan muchas dudas, como se vio en la experiencia de los EUA y la nuestra misma, y son muy costosos. Se recomienda aplicar métodos más sencillos, al menos en programas pequeños, y hacer experimentación sólo a nivel de la estrategia global o de grandes programas selectos.
- No se puede evaluar el impacto anualmente, no tiene sentido porque es muy poco tiempo para percibir diferencias claras provocadas por el programa y resulta muy costoso. Además los evaluadores en realidad no lo hacen, como descuidan otros aspectos solicitados por los términos de referencia (cobertura y su evolución, satisfacción de los usuarios, diseño de escenarios) o no solicitados pero importantes (participación social, utilización de resultados y recomendaciones previas).

- Resulta fundamental el respeto a un principio deontológico básico, que no se está respetando actualmente: la autonomía financiera e institucional del evaluador para garantizar su objetividad.
- Es imprescindible que una evaluación valore la gestión del programa. En los trabajos revisados hay datos, que en muchos casos se describen pero no se interpretan, hay pocas valoraciones explícitas y, sin embargo, todas hacen recomendaciones.
- Los diferentes actores entrevistados en esta metaevaluación reconocieron que el avance de la evaluación en México aún es limitado e identificaron como las principales causas de esta situación las culturales, sociales y políticas; en cambio, dieron un peso mucho menor a los problemas metodológicos que se relacionan en gran parte con la calidad de las evaluaciones analizadas.
- Aún no se cuenta con un avance suficiente de la profesionalización de la evaluación, entendida como la creación de una comunidad de expertos, que cuente con una revisión propia del conocimiento alcanzado, con base en criterios de calidad aceptados y el desarrollo de foros de discusión especializados.
- Tampoco puede decirse que se haya logrado la institucionalización de los procesos evaluativos a nivel gubernamental, porque aún no se han creado las organizaciones dedicadas exclusivamente al tema y, mucho menos existe un conjunto de reglas de juego (normas, rutinas, tradiciones, convenciones, etcétera), admitidas como una necesidad compartida por los diversos agentes y actores vinculados a ella.
- La inclusión de procesos de evaluación en el país no constituye una verdadera reforma del Estado si no se incorpora una gestión por resultados (ABR). En 2002 se estaba volviendo un mero rito, incrementaba el trabajo administrativo, creaba un nuevo mercado de trabajo para especialistas, pero no se integraba al proceso decisorio, por lo que incrementaba los costos y no justificaba su relación Costo/Beneficio.<sup>120</sup>

En resumen, se identifican dos grandes temas principales a encarar en torno a la función de evaluación en nuestro contexto actual, que recuerdan ampliamente la descripción realizada por Nioche (1993) para el contexto francés: su organización y metodología.

<sup>120</sup> En 2004 aún no se evaluaba la utilización de resultados inducida por Sedesol-SFP.

## Organización

Es necesario continuar la integración de programas evitando duplicidades, no sólo al interior de la Sedesol sino con otras secretarías que manejan programas similares,<sup>121</sup> reducir el número de instancias participantes en la evaluación y agilizar su funcionamiento, precisar la redacción de los términos de referencia, mejorar las relaciones interorganizacionales, eliminar prácticas estereotipadas, mejorar los sistemas de información, aumentar la legitimidad del proceso, organizar la participación social, desarrollar una verdadera cultura de la evaluación, utilizar sus resultados, etcétera.

## Metodología

Se requiere mejorar la calidad del muestreo aplicado y el análisis de la información recogida, evaluar la pertinencia del programa y su diseño, analizar la evolución del gasto ejercido en relación con las metas alcanzadas y su distribución geográfica, medir la cobertura como porcentaje de la población que necesita el servicio, medir realmente el impacto en los programas que lo sigan exigiendo, analizar el nivel de participación social efectiva, diseñar los escenarios prospectivos que se soliciten, analizar la utilización realizada de las recomendaciones de evaluaciones previas y, fundamentalmente, emitir un juicio valorativo final, fundamentado en todos los elementos revisados.

Lo anterior nos indica que los resultados alcanzados en materia de evaluación dependen de la eficacia interna de los procesos aplicados y de las condiciones organizacionales imperantes en su contexto. La similitud con el caso francés de inicio de los noventa, y las décadas de trabajos que nos separa de él, nos permiten aprender de su proceso, centrado especialmente en la solución de los problemas metodológicos (búsqueda de un enfoque alternativo al experimental), lo que lo condujo a postergar muchos de sus problemas organizacionales, que se han vuelto los más urgentes en la actualidad (Thoenig, 2002).

En síntesis, se concluye que el importante desarrollo alcanzado en los últimos años en materia de evaluación de políticas y programas públicos en México aún presenta problemas organizacionales, de calidad técnico-metodológica y de participación social.

Se considera que las razones que explican la situación anterior incluyen: presupuestos inadecuados; información institucional insuficiente; formación y experiencia limitada de los evaluadores, funcionarios y congresistas; falta

<sup>120</sup> Ejemplo: apoyo a la pobreza extrema entre los campesinos por la Sagarpa, a jornaleros agrícolas migrantes por la Secretaría de Trabajo y Previsión Social (STPS) o a jornaleros agrícolas por Sedesol.

de operacionalización de los mecanismos planeados de participación social; y, restricciones a la independencia de criterio de los evaluadores, a la transparencia de los procesos y la rendición de cuentas a la ciudadanía.

Finalmente, las propuestas formuladas para coadyuvar a la resolución de los problemas identificados en el trabajo han sido planteadas con la intención de que México aproveche la enorme oportunidad que tiene actualmente de mejorar la calidad de los trabajos de evaluación e integrar sus resultados a la gestión pública, como forma de incidir más efectivamente en el alivio del más urgente problema nacional: la pobreza en que vive la mitad de sus ciudadanos.

## Consideraciones finales

A MANERA de cierre de la revisión teórica, metodológica y empírica presentada en la primera parte sobre el tema de evaluación de los programas públicos, y de la metaevaluación aplicada en el segunda parte a las evaluaciones realizadas a los programas de desarrollo social en México, se considera relevante mencionar las nuevas tendencias del debate en la materia.

### Las tendencias en el debate teórico sobre evaluación de programas

Existe consenso sobre la necesidad de transformar a la administración pública tradicional para alcanzar de mejor manera los objetivos de las políticas y programas gubernamentales, pero no lo hay en cuanto a la dirección de dicho cambio, en el que se enfrentan quienes priorizan la solución de problemas técnicos y macroeconómicos apoyados en el NPM y quienes anteponen la solución de los conflictos y demandas sociales mediante el diálogo continuo entre gobierno y sociedad.

Esta disyuntiva se refleja en el terreno de la evaluación de programas sociales, también con dos tendencias claras: la aplicación de métodos experimentales y mediciones econométricas de trasfondo positivista, realizadas por especialistas; o, la utilización de métodos cualitativos que enfatizan la participación y el aprendizaje social, fundamentada en el constructivismo.

Si bien la literatura revisada muestra el énfasis experimental que la evaluación adquirió en EUA a finales de los sesenta y principios de los setenta, todo parece indicar que su utilización ha disminuido sensiblemente debido a sus altos costos y discutibles resultados, lo que coincide con los comentarios críticos que el BM dirige al método, por lo que sorprende que Sedesol siga insistiendo tanto en este enfoque. En cambio, la revisión hemerográfica en aquel país muestra que los estudios de carácter cualitativo están siendo más

productivos y aceptados por los lectores. Otro tanto sucede con el giro tomado por la evaluación en Francia.

En el caso de México se recomienda evitar posiciones dogmáticas, combinando metodologías según las necesidades de cada programa.

En relación con el componente de participación social, mientras México sigue cuestionándose sobre su conveniencia, en diversos países constituye una estrategia ampliamente aceptada (países escandinavos, Suiza, Francia, Canadá, etcétera), siendo su preocupación actual, la forma de mejorar la calidad de la misma. Por ello se insiste en la necesidad de poner a prueba su utilidad en nuestro contexto, aplicando una estrategia incremental. Su ausencia podría considerarse un síntoma de que en realidad el gobierno no se está preocupando por impulsar el aprendizaje democrático directo, por contar con ciudadanos informados, críticos y propositivos, sino sólo por utilizar las evaluaciones para racionalizar el gasto social y ganar legitimidad.

Por otra parte, si consideramos que desde finales de los sesenta Scriven ya trabajaba en la realización de metaevaluaciones en EUA, tenemos que reconocer un avance muy lento del tema, que en los últimos años comienza a parecer con más regularidad y en contextos nuevos (Canadá, Francia, España, Costa Rica), en la medida en que también se extiende el desarrollo de la evaluación de programas, su objeto de estudio. Además, comienzan a diferenciarse enfoques complementarios de metaevaluación: metodológico, hermenéutico, político, deontológico y organizativo.

Una tendencia relevante, detectada en todas las metaevaluaciones revisadas, es que los resultados de las evaluaciones son muy poco usados en los procesos decisorios, por lo que aún no es posible hablar de una gestión por resultados (Rist) o ABR (BM). Se espera que esta situación vaya cambiando a medida que se modifica la cultura evaluativa imperante en la administración pública y que se mejora la calidad de las evaluaciones practicadas.

### El reciente informe del Banco Mundial sobre la pobreza y las estrategias del gobierno mexicano

El BM (2004a) acaba de presentar un informe que confirma algunos problemas y soluciones expuestos en este libro. Reconoce el esfuerzo realizado por el gobierno mexicano, y especialmente Sedesol, en materia de monitoreo y evaluación de los programas dirigidos a la lucha contra la pobreza.<sup>122</sup> Pero,

<sup>122</sup> La reducción de la pobreza por ingresos mostraría resultados estadísticamente significativos a nivel rural y urbano; en cambio, la pobreza estimada en términos de consumo sólo los exhibiría en el ámbito rural.

identifica una agenda pendiente en materia social: integración más eficaz de la política económica y la social, calidad de los servicios, acceso a seguridad social, políticas para el incremento de ingresos y la mejora de su distribución (especialmente en el caso de grupos indígenas), mejora en la productividad de los autoempleados y de las empresas informales, fortalecimiento de la rendición de cuentas, la participación social y la administración basada en resultados (ABR), a partir de un enfoque integrado de M&E, como el que está financiando en Argentina.<sup>123</sup>

Opina que la estrategia Contigo sigue siendo difícil de evaluar y que no queda clara la relación entre su gran número de programas y el enfoque integral perseguido por el gobierno (variaciones en la focalización de acuerdo a la secretaría que lo implementa, por ejemplo). Agrega que se sabe poco respecto del gasto descentralizado en combate a la pobreza y que no queda claro el incentivo que induciría a los estados a interesarse en su aplicación.

Considera excelente la evaluación de impactos iniciada a Progresá en 1998, en la que pone más atención que en la actual evaluación de Oportunidades. Sin embargo, opina que se requiere comprender mejor los determinantes de la cobertura y la calidad de los servicios públicos dirigidos a la población pobre, así como a las condiciones institucionales que afectan a la incorporación social y la rendición de cuentas. El diseño experimental, si bien lo considera la técnica más novedosa de medición de impacto, opina que es "...más limitada para el aprendizaje continuo de los procesos que están detrás de los cambios observados y para comprender el papel del contexto local", así como "no puede controlar la influencia del tiempo en los resultados que no se relacionan con el programa" (BM, 2004a: lvii y 242), los que resultan decisivos para mejorar el diseño de los programas. Piensa que son los sistemas de monitoreo con base en el seguimiento continuo de indicadores los que permiten mayor aprendizaje y participación social, razón por la que propone su combinación con las evaluaciones en profundidad, tema en el que Sedesol trabaja actualmente.

También concuerda con las conclusiones de este trabajo en el sentido de que los programas requieren tiempo para que se evalúen sus impactos, lo que no puede hacerse todos los años, sino en forma selectiva, razón por la que considera que su propia evaluación debe considerarse como preliminar. Menciona que la realización de tantas evaluaciones de impacto lleva a que éstas sean rápidas, superficiales y arrojen pocos beneficios, por lo que recomienda su complementación con un sistema de monitoreo

<sup>123</sup> También en Sedesol se está proponiendo que la dirección encargada actualmente de la evaluación pase a llamarse Dirección General de Monitoreo y Evaluación de Programas Sociales.

como el que se estaría preparando en la oficina de la presidencia que permite una cierta anticipación de impactos. Además, sugiere interacciones entre evaluadores internos y externos, con énfasis en la participación de los beneficiarios. Considera que sólo tiene sentido valorar la efectividad de los programas sociales que sean innovadores y reproducibles, y tengan una alta cobertura nacional, intervenciones bien definidas y absorban recursos sustanciales (Banco Mundial, 2004a: 255).

El Banco opina que la LGDS enfrenta desafíos para su implementación debido a la rigidez que implica en material presupuestal, las funciones diseñadas para las comisiones creadas y el énfasis puesto en la gestión de los municipios.

Concluye que: 1) la Sedesol debe expandir en forma gradual pero continua su sistema de monitoreo y evaluación, tomando más en cuenta las necesidades de información de la propia secretaría, 2) el sistema requiere claridad de metas y un alineamiento de todos los funcionarios trabajando en ese sentido, lo que el banco no percibe actualmente, 3) es necesario redefinir el mandato del Congreso para no exigir evaluaciones anuales de impacto, sino un monitoreo del desempeño, y recurrir a una mezcla de evaluaciones internas y externas.

Todo lo anterior considera que debe complementarse con un mayor trabajo en temas de liderazgo, en el diseño de incentivos (sanciones y recompensas) y en la definición de funciones y responsabilidades.

### Expectativas de contribución al debate actual

El trabajo que finaliza aspira modestamente a realizar una contribución en torno a la evaluación en nuestro contexto en los siguientes aspectos:

- coadyuvar al conocimiento de enfoques y experiencias diferentes a las aplicadas en México para enriquecer el debate teórico y empírico;
- aportar algunas aclaraciones conceptuales y metodológicas;
- identificar y relacionar problemas en el proceso de evaluación, así como realizar sugerencias para su solución;
- ofrecer formas de organización alternativas para la coordinación de los procesos evaluativos;
- insistir en las particularidades de los programas sociales y sus evaluaciones;



- proponer su evaluación mediante métodos cuantitativos y cualitativos integrados, y simplificar los actuales en el caso de programas pequeños;
- complementar la metaevaluación basada en la revisión del informe final y los problemas metodológicos con el análisis de su entorno organizacional; y
- ofrecer una propuesta de métodos y criterios para la realización de metaevaluaciones en nuestro contexto.

Finalmente, si bien las metaevaluaciones se encuentran aún muy poco difundidas y no cuentan con métodos específicos desarrollados, se espera haber contribuido a mostrar su utilidad y se recomienda su realización periódica para ayudar a los tomadores de decisiones a la reflexión comprometida pero independiente sobre la forma en que se están aplicando las evaluaciones en curso y a la generación de un proceso de acumulación de conocimientos y aprendizaje social.

Anexo 1

Normas para la evaluación de programas  
del Joint Committee on Standards  
for Educational Evaluation  
Segunda edición, 1994

Normas de utilidad

Las normas de utilidad pretenden asegurar que una evaluación servirá a las necesidades de información de los usuarios pretendidos.

*U1 Identificación de los agentes críticos (stakeholders).* Se deben identificar las personas implicadas o afectadas por la evaluación, de manera que sus necesidades puedan ser cubiertas.

*U2 Credibilidad del evaluador.* Las personas que dirijan una evaluación deben ser merecedoras de confianza y competentes para realizar una evaluación, de manera que los hallazgos de la evaluación consigan la máxima credibilidad y aceptación.

*U3 Selección y alcance de la información.* La información recopilada debe ser seleccionada ampliamente para contestar preguntas pertinentes sobre el programa y debe responder a las necesidades e intereses de clientes y otros agentes críticos.

*U4 Identificación de valores.* Las perspectivas, los procedimientos y la fundamentación que se utilicen al interpretar los resultados deben describirse con cuidado, de manera que los fundamentos de los juicios de valor sean claros.

*U5 Claridad del informe.* Los informes de evaluación deben describir claramente el programa que está siendo evaluado, incluyendo su contexto, y los propósitos, procedimientos y resultados de la evaluación, de manera que se ofrezca información esencial y fácilmente comprensible.

*U6 Oportunidad y difusión del informe.* Los hallazgos intermedios significativos y los informes de evaluación se deben difundir a los usuarios deseados, de manera que puedan ser utilizados de manera oportuna.

*U7 Impacto de la evaluación.* Las evaluaciones se deben planificar, dirigir y comunicar de forma que animen a los agentes críticos a su cumplimiento, de manera que aumente la probabilidad de que se use la evaluación.

## Normas de factibilidad/viabilidad

Las normas de factibilidad tienen el propósito de asegurar que una evaluación sea realista, prudente, diplomática y frugal.

*F1 Procedimientos prácticos.* Los procedimientos de la evaluación deben ser prácticos, de manera que las interrupciones sean mínimas mientras se obtiene la información necesaria.

*F2 Viabilidad política.* La evaluación se debe planificar y dirigir anticipando las diferentes posturas de diversos grupos de interés, de manera que se pueda obtener su colaboración y así se puedan advertir y contrarrestar posibles intentos de restringir las operaciones de evaluación o de sesgar o abusar de los resultados por parte de estos grupos.

*F3 Coste efectividad.* La evaluación debe ser eficiente y producir una información de valor suficiente, de manera que se justifiquen los recursos empleados.

## Normas de corrección/legitimidad

Son las normas dirigidas a asegurar que la evaluación se dirija legal y éticamente y con la debida atención al bienestar de aquellos implicados en la evaluación, así como a aquellos afectados por sus resultados.

*C1 Orientación de servicio.* Las evaluaciones se deben diseñar para asistir a las organizaciones a tratar y cubrir efectivamente las necesidades de todo el espectro de destinatarios.

*C2 Acuerdos formales.* Las obligaciones de las partes formales de la evaluación (qué se hará, cómo, por quién y cuándo) se deben acordar por escrito, de forma que dichas partes estén obligadas a seguir todas las condiciones del acuerdo o, por el contrario, renegociarlo formalmente.

*C3 Derechos de los sujetos.* Las evaluaciones se deben diseñar y dirigir de tal forma que se respeten y protejan los derechos y el bienestar de los sujetos.

*C4 Interacciones humanas.* Los evaluadores deben respetar la dignidad y el valor humano en sus interacciones con otras personas asociadas a la evaluación, de manera que no se amenace o dañe a los participantes.

*C5 Valoración completa y justa.* La evaluación debe ser completa y justa en su análisis y en su registro de fortalezas y debilidades del programa evaluado, de tal modo que las fortalezas sean reforzadas y las áreas problemáticas tratadas.

*C6 Revelación de los resultados.* Las partes formales de la evaluación deben asegurar que el conjunto completo de resultados de la evaluación junto con sus pertinentes limitaciones son accesibles para las personas afectadas por la evaluación y para cualquier otra persona con derechos legales para recibir los resultados.

*C7 Conflicto de intereses.* Los conflictos de intereses deben manejarse abierta y honestamente de tal forma que no comprometan el proceso ni los resultados de la evaluación.

*C8 Responsabilidad fiscal.* La distribución y los gastos de recursos asignados por el evaluador deben reflejar procedimientos solventes de información y responsabilidad, en caso contrario el evaluador debe ser prudente y éticamente responsable de manera que los gastos se justifiquen y sean apropiados.

## Normas de precisión

Estas normas tienen el propósito de asegurar que una evaluación revele y proporcione información técnicamente apropiada sobre las características que determinan el valor o mérito del programa evaluado.

*P1 Documentación del programa.* El programa evaluado debe ser descrito y documentado de forma apropiada y clara, de tal manera que el programa se identifique claramente.

*P2 Análisis del contexto.* El contexto en el cual el programa existe se debe examinar con suficiente detalle de manera que se puedan identificar sus posibles influencias sobre dicho programa.

*P3 Descripción de propósitos y procedimientos.* Los propósitos y procedimientos de la evaluación deben ser seguidos y descritos con el suficiente detalle de manera que se puedan identificar y valorar.

*P4 Fuentes de información defendibles.* Las fuentes de información que se utilicen en una evaluación de programa se deben describir con suficiente detalle de forma que se pueda estimar la adecuación de la información.

*P5 Información válida.* Los procedimientos de recopilación de información se deben elegir y desarrollar y entonces poner en práctica de manera que se asegure que la interpretación a la que se ha llegado es válida para el uso que se le pretende dar.

*P6 Información fiable.* Los procedimientos de recopilación de información se deben elegir, desarrollar y poner en práctica de manera que se asegure que la información obtenida es suficientemente fiable para el uso que se le pretende dar.

*P7 Información sistemática.* La información recopilada, procesada y transmitida en una evaluación debe ser revisada sistemáticamente y cualquier error encontrado debe ser corregido.

*P8 Análisis de la información cuantitativa.* La información cuantitativa de una evaluación se debe analizar apropiada y sistemáticamente de manera que las preguntas de evaluación se contesten de forma efectiva.

*P9 Análisis de la información cualitativa.* La información cualitativa de una evaluación se debe analizar apropiada y sistemáticamente de manera que las preguntas de evaluación se contesten de forma efectiva.

*P10 Conclusiones justificadas.* Las conclusiones a las que se llega en una evaluación deben estar justificadas explícitamente de manera que los agentes críticos de la evaluación puedan ponderarlas.

*P11 Informe imparcial.* Los procedimientos para informar deben asegurar que no haya distorsión causada por sesgos y sentimientos personales de cualquiera de las partes de la evaluación, de manera que los informes reflejen de forma justa los resultados de la evaluación.

*P12 Metaevaluación.* La evaluación en sí misma se debe evaluar formativa y sumativamente a través de éstas y otras normas pertinentes, de manera que se guíe de forma apropiada su realización y, en cuanto se termine, los agentes críticos puedan examinar de cerca sus fortalezas y debilidades.

## Anexo 2

# Principales funciones de las secretarías de Estado vinculadas a los procesos de evaluación

### Secretaría de Hacienda y Crédito Público (SHCP)

A CONTINUACIÓN se señalan sus funciones vinculadas a la evaluación, sus presupuestos y la utilización de sus resultados:

Inciso XIV- Proyectar y calcular los egresos del Gobierno Federal y de la administración pública paraestatal, haciéndolos compatibles con la disponibilidad de recursos y en atención a las necesidades y políticas del desarrollo nacional.

Inciso XV- Formular el programa del gasto público federal y el proyecto de Presupuesto de Egresos de la Federación y presentarlos, junto con el del Departamento del Distrito Federal (*actualmente Gobierno del Distrito Federal*), a la consideración del Presidente de la República.

Inciso XVI- Evaluar y autorizar los programas de inversión pública de las dependencias y entidades de la administración pública federal.

Inciso XVII- Llevar a cabo las tramitaciones y registros que requiera la vigilancia y evaluación del ejercicio del gasto público federal y de los presupuestos de egresos.

Inciso XXIII- Vigilar el cumplimiento de las obligaciones derivadas de las disposiciones en materia de planeación nacional, así como de programación, presupuestación, contabilidad y evaluación” (Ley Orgánica de la administración pública federal, artículo 31).

### Secretaría de la Función Pública (SFP)

A continuación se señalan sus funciones vinculadas a nuestro objeto de estudio:

Inciso I- Organizar y coordinar el sistema de control y evaluación gubernamental. Inspeccionar el ejercicio del gasto público federal, y su congruencia con los presupuestos de egresos.

Inciso II- Expedir las normas que regulen los instrumentos y procedimientos de control de la administración pública federal, para lo cual podrá requerir de las dependencias competentes, la expedición de normas complementarias para el ejercicio del control administrativo.

Inciso III- Vigilar el cumplimiento de las normas de control y fiscalización así como asesorar y apoyar a los órganos de control interno de las dependencias y entidades de la administración pública federal.

Inciso IV- Establecer las bases generales para la realización de auditorías en las dependencias y entidades de la administración pública federal, así como realizar las auditorías que se requieran a las dependencias y entidades en sustitución o apoyo de sus propios órganos de control.

Inciso V- Vigilar el cumplimiento, por parte de las dependencias y entidades de la administración pública federal, de las disposiciones en materia de planeación, presupuestación, ingresos, financiamiento, inversión, deuda, patrimonio, fondos y valores.

Inciso VII- Realizar, por sí o a solicitud de la Secretaría de Hacienda y Crédito Público o de la coordinadora del sector correspondiente, auditorías y evaluaciones a las dependencias y entidades de la administración pública federal, con el objeto de promover la eficiencia en su gestión y propiciar el cumplimiento de los objetivos contenidos en sus programas.

Inciso VIII- Inspeccionar y vigilar, directamente o a través de los órganos de control, que las dependencias y entidades de la administración pública federal cumplan con las normas y disposiciones en materia de: sistemas de registro y contabilidad, contratación y remuneraciones de personal, contratación de adquisiciones, arrendamientos, servicios, y ejecución de obra pública, conservación, uso, destino, afectación, enajenación y baja de bienes muebles e inmuebles, almacenes y demás activos y recursos materiales de la administración pública federal.

Inciso X- Designar a los auditores externos de las entidades, así como normar y controlar su desempeño. Designar, para el mejor desarrollo del sistema de control y evaluación gubernamentales, delegados de la propia Secretaría ante las dependencias y órganos desconcentrados de la administración pública federal centralizada, y comisarios en los órganos de gobierno o vigilancia de las entidades de la administración pública paraestatal.

Inciso XI- Designar y remover a los titulares de los órganos internos de control de las dependencias y entidades de la administración pública federal y de la Procuraduría General de la República, así como a los de las áreas de auditoría, quejas y responsabilidades de tales órganos, quienes dependerán jerárquica y funcionalmente de la Secretaría de Contraloría y Desarrollo

Administrativo (*actualmente Secretaría de la Función Pública*), tendrán el carácter de autoridad y realizarán la defensa jurídica de las resoluciones que emitan en la esfera administrativa y ante los Tribunales Federales, representando al Titular de dicha Secretaría.

Inciso XII- Colaborar con la Contaduría Mayor de Hacienda (*actualmente Auditoría Superior de la Federación*) para el establecimiento de los procedimientos necesarios que permitan a ambos órganos el mejor cumplimiento de sus respectivas responsabilidades.

Inciso XIII- Informar periódicamente al Ejecutivo federal, sobre el resultado de la evaluación respecto de la gestión de las dependencias y entidades de la administración pública federal, así como de aquellas que hayan sido objeto de fiscalización, e informar a las autoridades competentes, cuando proceda del resultado de tales intervenciones y, en su caso, dictar las acciones que deban desarrollarse para corregir las irregularidades detectadas.

Inciso XXV- Formular y conducir la política general de la administración pública federal para establecer acciones que propicien la transparencia en la gestión pública, la rendición de cuentas y el acceso por parte de los particulares a la información que aquélla genere.” (Ley Orgánica de la administración pública federal, artículo 37).



## Siglas y acrónimos

- AECA: Asociación Española de Contabilidad y Administración de Empresas.
- AL: América Latina.
- ASF: Auditoría Superior de la Federación.
- BID: Banco Interamericano de Desarrollo.
- BM: Banco Mundial.
- Canacindra: Cámara Nacional de la Industria de la Transformación.
- CAP: Comisión de Administración Pública.
- C/B: Costo/Beneficio.
- CCE: Consejo Científico de Evaluación.
- C/E: Costo/Eficacia.
- CEPAL: Comisión Económica para América Latina.
- CERTU: Centre d'Études sur les Réseaux de Transport et l'Urbanisme.
- CIDE: Centro de Investigación y Docencia Económicas.
- CIESAS: Centro de Investigación y Estudios Superiores de Antropología Social.
- CIME: Comité Interministerial de Evaluación.
- CMH: Contaduría Mayor de Hacienda.
- CNDH: Comisión Nacional de Derechos Humanos.
- CNDPI: Comisión Nacional de Desarrollo de los Pueblos Indígenas.
- CNE: Consejo Nacional de Evaluación de Políticas Públicas (Francia).
- CNE: Consejo Nacional de Evaluación de la Política de Desarrollo Social (México).
- Colmex: Colegio de México.
- Conacyt: Consejo Nacional de Ciencia y Tecnología.
- Conapo: Consejo Nacional de Población.
- Conasupo: Comisión Nacional de Subsistencias Populares.
- Conaza: Comisión Nacional de Zonas Áridas.
- Conprograsa: Coordinación Nacional de Prograsa.

- Coplamar: Coordinación del Plan General de Zonas Deprimidas y Grupos Marginados.
- CPCP: Comisión de Presupuesto y Cuenta Pública.
- Diconsa: Distribuidora Conasupo.
- DIF: Desarrollo Integral de la Familia.
- EES: European Evaluation Society.
- ELECTRA: Élection et choix traduisant la réalité.
- ENAH: Escuela Nacional de Antropología e Historia.
- ENIGH: Encuesta Nacional de Ingresos y Gastos de los Hogares.
- Encel: Encuesta de Evaluación en Localidades.
- EUA: Estados Unidos de América.
- E&M: Evaluación y Monitoreo.
- FAO: Organización de las Naciones Unidas para la Agricultura y la Alimentación.
- FCPYS: Facultad de Ciencias Políticas y Sociales.
- Fidelist: Fideicomiso para la Liquidación al Subsidio de la Tortilla.
- FMI: Fondo Monetario Internacional.
- FNDE: Fondo Nacional para el Desarrollo de la Evaluación.
- Fonaes: Fondo Nacional de Empresas en Solidaridad.
- Fonart: Fondo Nacional de las Artesanías.
- Fonhapo: Fondo Nacional de las Habitaciones Populares.
- Funsalud: Fundación Mexicana para la Salud.
- GAO: General Accounting Office.
- GAPP: Grupo de Análisis de Políticas Públicas.
- GEA: Grupo de Economistas Asociados.
- GIES: Grupo Interdisciplinario de Expertos en Investigaciones Sociales y Políticas.
- IAWG: Inter-Agency Working Group on Evaluation.
- IFAI: Instituto Federal de Acceso a la Información Pública.
- IFPRI: Internacional Food Policy Research Institute.
- IMSS: Instituto Mexicano del Seguro Social.
- INAH: Instituto Nacional de Antropología e Historia.
- Inapam: Instituto Nacional de Personas Adultas Mayores.
- IndesoL: Instituto Nacional de Desarrollo Social.
- Indetec: Instituto para el Desarrollo Técnico de las Haciendas Públicas.
- INEGI: Instituto Nacional de Estadística, Geografía e Informática.
- Infonavit: Instituto Nacional del Fondo para la Vivienda de los Trabajadores.
- INI: Instituto Nacional Indigenista.

- INSP: Instituto Nacional de Salud Pública.
- ISSSTE: Instituto de Seguridad y Servicios Sociales para los Trabajadores del Estado.
- ITAM: Instituto Tecnológico Autónomo de México.
- ITESM: Instituto Tecnológico de Estudios Superiores de Monterrey.
- LGDS: Ley General de Desarrollo Social.
- Liconsa: Leche Industrializada Conasupo.
- MCE: Misión de Evaluación y Control del Gasto Público.
- Mideplan: Ministerio de Planificación Nacional y Política Económica.
- MIT: Instituto Tecnológico de Massachusett.
- NAO: National Audit Office.
- NPM: New Public Mangement.
- OCDE: Organización para la Cooperación y el Desarrollo Económico.
- OEA: Organización de los Estados Americanos.
- OIT: Organización Internacional del Trabajo.
- OMB: Office of Management and Budget.
- OMS: Organización Mundial de la Salud.
- ONG: Organización No Gubernamental.
- ONU: Organización de las Naciones Unidas.
- OSC: Organizaciones de la Sociedad Civil.
- PAC: Programa de Ampliación de Cobertura en Salud.
- PAN: Partido Acción Nacional.
- PEF: Presupuesto de Egresos de la Federación.
- PHSD: Programa para los Habitantes del Semi Desierto.
- PIDER: Programa de Inversiones Públicas para el Desarrollo Rural.
- PND: Plan Nacional de Desarrollo.
- PNDS: Programa Nacional de Desarrollo Social.
- PMG: Programas de Mejoramiento de Gestión.
- PNUD: Programa de Naciones Unidas para el Desarrollo.
- PNUMA: Programa de Naciones Unidas para el Medio Ambiente.
- PPBS: Planning Programming Budgeting System.
- PRD: Partido de la Revolución Democrática.
- PRI: Partido Revolucionario Institucional.
- Prodim: Programa de Desarrollo Institucional Municipal.
- Progresa: Programa de Educación, Salud y Alimentación.
- Promap: Programa para la Modernización de la Administración Pública Federal.
- Pronasol: Programa Nacional de Solidaridad.
- PRSP: Poverty Reduction Strategy Paper.

- RCB: Rationalisation des Choix Budgetaires.
- Sagarpa: Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación.
- Sahop: Secretaría de Asentamientos Humanos y Obras Públicas.
- SAM: Sistema Alimentario Mexicano.
- SARH: Secretaría de Agricultura y Recursos Hidráulicos.
- Secodam: Secretaría de la Contraloría y el Desarrollo Administrativo.
- Secogef: Secretaría General de la Federación.
- SCT: Secretaría de Comunicaciones y Transporte.
- Sedesol: Secretaría de Desarrollo Social.
- SEGPR: Sistema de Evaluación de la Gestión Pública por Resultados.
- Semarnat: Secretaría de Medio Ambiente y Recursos Naturales.
- SEP: Secretaría de Educación Pública.
- SEV: Sistema Informático de Evaluación de Gestión Pública.
- SFE: Sociedad Francesa de Evaluación.
- SFP: Secretaría de la Función Pública.
- SHCP: Secretaría de Hacienda y Crédito Público.
- SIDA: Síndrome de Inmuno Deficiencia Adquirida.
- SIGFE: Sistema de Información para la Gestión Financiera del Estado.
- Sine: Sistema Nacional de Evaluación.
- Sinergia: Sistema Nacional de Evaluación de Resultados de la Gestión.
- Sisep: Sistema de Evaluación y Seguimiento de Proyectos.
- SNPP: Sistema Nacional de Planeación Participativa.
- SPP: Secretaría de Programación y Presupuesto.
- SSA: Secretaría de Salud.
- STPS: Secretaría de Trabajo y Previsión Social.
- TIR: Tasa interna de retorno.
- UIA: Universidad Iberoamericana.
- NU: Naciones Unidas.
- UAM-A: Universidad Autónoma Metropolitana, plantel Azcapotzalco.
- UAM-X: Universidad Autónoma Metropolitana, plantel Xochimilco.
- UNAM: Universidad Nacional Autónoma de México.
- UNESCO: Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura.
- VAN: Valor Actual Neto.

## Bibliografía

### Libros y revistas

- ABATE, Bernard, 2000, *La nouvelle gestion publique*, Collection Systemès, París, L.G.D.J.
- AGUILAR, María José y Ezequiel Ander-Egg, 1992, *Evaluación de servicios y programas sociales*, España, Siglo XXI.
- , 1998, *Cómo elaborar un proyecto. Guía para diseñar proyectos sociales y culturales*, Colección Política, Servicios y Trabajo Social, Buenos Aires, Lumen–Humanitas.
- AGUILAR, Luis F, compilación y estudio introductorio, 2000a, *El estudio de las políticas públicas*, 3a. ed., México, M.A. Porrúa.
- , compilación y estudio introductorio, 2000b, *Problemas públicos y agenda de gobierno*, 3a ed., México, M.A. Porrúa.
- , compilación y estudio introductorio, 2000c, *La hechura de las políticas*, 3a ed., México, M.A. Porrúa.
- , 2000d, “La reforma del Estado mexicano”, en José Luis Méndez (comp.), *Lecturas básicas de administración y políticas públicas*, México, Colmex.
- ALLISON, Graham, 1971, *Essence of decision. Explaining of Cuban missile crisis*, Boston, Little, Brown and Co.
- ANDER-EGG, Ezequiel, 1984, *Evaluación de programas de trabajo social*, Buenos Aires, Humanitas.
- ANDERSON, 2000, *Public policy making*, 4a. ed., Boston, Houghton Mifflin Co.
- BALLART, Xavier, 1992, *¿Cómo evaluar programas y servicios públicos? Aproximación sistemática y estudios de caso*, Madrid, Ministerio para las Administraciones Públicas.
- BANCO MUNDIAL, 2001, *Attacking poverty*, World development report 2000/2001, Washington, D.C.
- , 2004, *La pobreza en México, una evaluación de las condiciones, las tendencias y la estrategia del gobierno*, México, Banco Mundial.

- BANCO INTERAMERICANO DE DESARROLLO, 2000, *Economic and social report 1998-1999*, Washington, D.C., IDB.
- BAÑÓN I MARTÍNEZ, Rafael, (comp.), 2003, *La evaluación de la acción y de las políticas públicas*, Madrid, Díaz de Santos.
- BARZELAY, Michael, 1992, *Breaking through bureaucracy*, California, University of California Press.
- , 2001, *The New Public Management. Improving research and policy dialogue*, Nueva York, University of California Press, Russell Sage Foundation.
- BELTRÁN, Miguel, 1991, *La realidad social*, Madrid, Tecnos.
- BEMELMANS-VIDEC, Marie Louise, Ray Rist y Evert Vedung (eds), 1998, *Carrots, sticks and sermons*, Colección Comparative Policy Analysis, EUA, Ray Rist editor.
- BERTHELOT, Jean Michel, 2001, *Épistémologie des sciences sociales*, París, PUF.
- BICHMANN, Rifkin y Shrestha, 1989, "Cómo medir la participación de la comunidad", *Foro Mundial de la Salud*, vol. 10.
- BIRDSALL, Nancy, D. Roo y R. Sabot, 1995, *La desigualdad como limitación del crecimiento en América Latina*, EUA, BID.
- BLANET, J. y et al., 2000, *Learning for change, issues and experiences in participatory monitoring and evaluation*, Londres, Intermediate Technology Publications.
- BOLTVINIK, Julio y E. Hernández, 1999, *Pobreza y distribución del ingreso en México*, México, Siglo XXI Editores.
- BOURDIEU, Pierre (dir.), 1993, *La misère du monde*, París, Ed. Du Seuil, Col. Libre Examen.
- BOUSSAGUET, Jacquot y Ravinet, 2004, *Dictionnaire des politiques publiques*, París, Presses de Sciences Po.
- BOYLE, Richard y Donald Lemaire (eds.), 1998, *Building effective evaluation capacity*, Colección Comparative Policy Analysis, EUA, Ray Rist editor.
- BOZEMAN, Barry, 1998, *La gestión pública. Su situación actual*, México, Colegio Nacional de Ciencias Políticas y Administración Pública-FCE.
- BRESSER PEREIRA, Luis et al., 2004, *Política y gestión pública*, Buenos Aires, FCE.
- BREWER, Garry y Peter de León, 1983, *The foundations of policy analysis*, EUA, Dorsey Press.
- BROWN, Deryck, 2001, "Evaluación de políticas y programas de la región del Caribe de habla inglesa, problemas conceptuales y prácticos", en *Serie Gestión Pública*, núm. 9, enero, Santiago de Chile, CEPAL-ILPES.

- BRUGUÉ, Quim, 2004, "Modernizar la administración desde la izquierda, burocracia, nueva gestión pública y administración deliberativa", en revista del CLAD, *Reforma y Democracia*, núm. 29, junio.
- BUSTELO, María, 2003, "¿Qué tiene de específico la metodología de la evaluación?", en Bañón (comp.), *La evaluación de la acción y de las políticas públicas*, Madrid, Díaz de Santos.
- , 2004, *La evaluación de políticas públicas de igualdad de género en España*, Madrid, Catarata.
- CAMPBELL, D. y J. Stanley, 1966, *Experimental and quasi-experimental designs for research*, Chicago, Rand McNally.
- CARDOZO BRUM, Myriam, 1986, "Exposición de métodos de evaluación no tradicionales", Serie Material Docente núm. 9, DAP, México, CIDE.
- , 1993, "La evaluación de las políticas públicas, problemas, metodologías, aportes y limitaciones", en revista de *Administración Pública* núm. 84, enero-junio, México, INAP.
- , 1998, "La función de evaluación en las organizaciones no lucrativas, una revisión de las preocupaciones conceptuales, teóricas y metodológicas recientes", revista *Administración y Organizaciones*, año 1, núm. 1, noviembre, México, UAM-X.
- , 1998, "Evaluación, conceptos, métodos y aplicaciones en educación", en Luis Berruecos (coord. y ed.), *La Evaluación en el Sistema Modular*, México, UAM-X.
- , 1999, "La evaluación en las actividades no lucrativas. Revisión de experiencias recientes en países desarrollados y situación en México", revista *Administración y Organizaciones*, año 1, núm. 2, julio, México, UAM-X.
- , 2000a, "La evaluación del desempeño", en J.L. Méndez y U. Raich, (coord.), *Evaluación del desempeño y servicio civil de carrera en la administración pública y los organismos electorales*, México, IFE-INAP.
- , 2000b, "Gerencia pública y participación social", en revista *Acta Sociológica*, núms. 28-29, enero-agosto, México, Facultad de Ciencias Políticas y Sociales, UNAM.
- , 2001, "Los métodos cuantitativos para el análisis de políticas públicas", en revista *Venezolana de Gerencia*, año 6, núm. 13, enero-marzo Venezuela, Universidad del Zulia.
- , 2003, "Evaluación de políticas de desarrollo social", en revista *Política y Cultura*, núm. 20, ¿Desarrollo Social?, otoño, México, UAM-X.
- CEPAL, 1998, "Gestión de programas sociales en América Latina", vol. I, *Serie Políticas Sociales* núm. 25, junio, Santiago de Chile.

- CEPAL, 2003, *Social panorama of Latin America 2002-2003*, Naciones Unidas.
- CIME-Primer Ministro-Comissariat Général du Plan, 1997, *La loi quinquennale relative au travail, à l'emploi et à la formation professionnelle*, Francia, la Documentation Française.
- CNE-Comissariat Général du Plan, 2002, *La politique de lutte contre le SIDA 1994-2000. Rapport de l'instance d'évaluation présidée par Christian Rollet*, Francia, la Documentation Française.
- COHEN, Ernesto, Rodrigo Martínez, Leonel Tapia y Alejandro Medina, 1998, "Metodología de análisis de la gestión de programas sociales", vol. II, *Serie Políticas Sociales*, núm. 25, junio Santiago de Chile, CEPAL.
- y R. Franco, 1993, *Evaluación de proyectos sociales*, México, Siglo XXI.
- , 2001, "Reforma del Estado, modernización administrativa y evaluación del desempeño de la gestión", en *Los desafíos de la reforma del Estado en los programas sociales, tres estudios de caso*, Serie Políticas Sociales núm. 45, Chile, CEPAL.
- COMMAILLES, Jacques y Bruno Jobert, (dirs), 1998, *Les métamorphoses de la régulation politique*, núm. 24, Francia, Droit et Societé.
- CONSEJO CIENTÍFICO DEL CLAD, 2000, *La responsabilización en la nueva gestión pública latinoamericana*, Venezuela, CLAD.
- COOKE y Kothary, 2001, *Participation, the new tyranny?*, Londres, Zed Books.
- CROZIER, Michel, 1974, *El fenómeno burocrático*, Buenos Aires, Amorrortu.
- , 1989, *Estado modesto, Estado moderno*, México, FCE.
- y J.C. Thoenig, 1975, "La régulation des systemes organisés complexes", en *Revue Française de Sociologie*, XVI, Francia.
- CUNILL, Nuria, 1995 "La rearticulación de las relaciones Estado-sociedad, en búsqueda de nuevos sentidos", en revista del CLAD, *Reforma y Democracia*, núm. 4.
- , 1997, *Repensando lo público a través de la sociedad, nuevas formas de gestión pública y representación social*, Venezuela, CLAD.
- y Sonia Ospina, 2003, *Evaluación de resultados para la gestión pública moderna y democrática, experiencias latinoamericanas*, Venezuela, CLAD-AECIMAP-FIIAPP.
- CHANUT, Véronique, 2002, "L'évaluation, affaire d'État ou question d'organisations?", en *Politiques et Management Public*, vol. 2, núm. 4, Número especial, Francia, Institut de Management Public.
- CHAUVIRÉ, C. y O. Fontaine, 2003, *Le vocabulaire de Bourdieu*, Francia, Ellipses.
- CHRISTENSEN, Tom y Per Laegreid, 2001, "La nueva administración pública, el equilibrio entre la gobernanza política y la autonomía administrativa", en revista *Gestión y Política Pública*, vol. X, núm. 1, México, CIDE.



- DEBONO, E., 1999, *Six thinking hats*, Londres, Backbay.
- DELEAU, M., J.P. Nioche, P. Penz y R. Poinard, 1986, *Evaluer les politiques publiques, méthodes, déontologie, organisation*, Francia, La Documentation Française.
- DEL VALLE, María Del Carmen y Diana Roldós, 2003, "Programa de leche, aciertos, limitaciones y perspectivas", en Torres Salcido (coord.), *Políticas de abasto alimentario. Alternativas para el Distrito Federal y su zona metropolitana*, México, ediciones Casa Juan Pablos, Facultad de Ciencias Políticas y Sociales, UNAM.
- DENTE, Bruno, 1998, "La politique de l'évaluation des politiques publiques", en Marie Christine Kesler, *et al.*, *Evaluation des politiques publiques*, París, L'Harmattan.
- DÍAZ, Luis Fernando, 2002, "La metaevaluación y su método", en *revista de Ciencias Sociales*, año XLIII, núms. 92-93, Costa Rica, Universidad de Costa Rica.
- , Rosay Rosales, 2003, *Metaevaluación, Evaluación de la evaluación de políticas, programas y proyectos sociales*, San José de Costa Rica, Editorial Universidad Estatal a Distancia.
- DROR, Yehezkel, 1970, "Prolegomena to policy sciences", en *Policy Science*, núm. 1, EUA.
- , 1977, "Aplicaciones del método de sistemas y técnicas cuantitativas en la formulación de políticas", en G. Campero, y H. Vidal (comp.), *Teoría general de sistemas y administración pública*, Costa Rica, Universidad Centroamericana.
- DURAN, Patrice y E. Monnier, 1992, "Le développement de l'évaluation en France. Nécessités techniques et exigences politiques", en *Revue Française de Science Politique*, vol. 40, núm. 2.
- , J.C. Thoenig, 1996, "L'État et la gestion territoriale", en *Revue Française de Science Politique*, agosto, Francia.
- , 1999, *Penser l'action publique*, Droit et Société, vol. 27, Francia.
- , 2002, "L'évaluation, au risque de son institutionalisation", en *Politiques et Management Public*, vol. 2, núm. 4, número especial, diciembre Francia, Institut de Management Public.
- DUVERGER, Maurice, 1983, *Sociología de la política*, Barcelona, Ariel.
- FÉRLIE, E. *et al.*, 1996, *The new public management in action*, Oxford, Oxford University Press.
- FERNÁNDEZ TORO, Julio César, 2004, "Presentación", en *NotiCLAD*, núm. 23, junio, Venezuela, CLAD.
- FINK, A. y J. Kosecoff, 1978, *An evaluation primer*, Washington, Washington Capitol Publications.

- FISHER, Frank, 1997, *Evaluating public policy*, Chicago, Nelson Hall Publishers.
- FRADE RUBIO, Laura (comp. y coord.), 2000, *¿Equidad, participación y correspondencia?*, México, Campaña El Banco Mundial en la Mira de las Mujeres.
- FREYSSINET, Jacques, 1998, “La fonction social de l'évaluation des dispositifs d'insertion professionnel des jeunes. Une comparaison européenne”, en Marie Christine Kesler *et al.*, *Evaluation des politiques publiques*, París, L'Harmattan.
- GAPP, 2003, *La lettre du GAPP*, núm. 0, Cachan, Francia, CNRS-GAPP-CNS.
- GLASS, Gene, 1976, “Primary, secondary and meta-analysis of research”, en *Educational Researcher*, núm. 5.
- GOLDSWORTHY, David, 1998, “Le controle du bon usage des fonds publics par le National Audit Office”, en Conseil Scientifique de l'Évaluation, *L'évaluation en développement 1997*, Francia, La Documentation Française.
- GORDON, Sara, 1997, “La cultura política de las ONG en México”, en revista *Mexicana de Sociología*, año LIX, núm. 1, enero-marzo, México, IIS-UNAM.
- , 1999, “Pobreza y programas de política social en México”, en revista *de Trabajo Social*, núm. 23, México, Escuela Nacional de Trabajo Social- UNAM.
- , 2000, “Nuevas desigualdades y política social”, en Cordera y Ziccardi (coords.), *Las políticas sociales de México al fin del milenio. Descentralización, diseño y gestión*, México, M.A. Porrúa-UNAM.
- , 2001, “Ciudadanía y derechos sociales, ¿criterios distributivos?”, en A. Ziccardi (comp.), *Pobreza, desigualdad y ciudadanía. Los límites de las políticas sociales en América Latina*, Buenos Aires, IIS, UNAM-CLACSO-FLACSO, México, ASDI.
- GUADARRAMA, Gloria, 2003, “Gerencia pública y política social en América Latina”, en revista *Economía, Sociedad y Territorio*, vol. IV, núm. 13, enero-junio, México, El Colegio Mexiquense A.C.
- GUBA, E.G., 1987, *Naturalistic evaluation. New direction for program evaluation*, núm. 34, San Francisco, Jossey Bass.
- , 1990, *The paradigm dialog*, Newsbury Park Publications.
- y V. Caracelli, 1989, *Fourth generation evaluation*, EUA, Sage.
- y Lincoln, 1981, *Effective evaluation*, San Francisco, Jossey Bass.
- GUERRERO, Juan Pablo, 1995, “La evaluación de políticas públicas, enfoques teóricos y realidades de nueve países desarrollados”, en revista *de Gestión y Política Pública*, núm. 1, vol. IV, primer semestre, México, CIDE, A.C.

- GUERRERO, Omar, 1998, "El management público, una torre de Babel", en revista *Convergencia*, año 5, núm. 17, México, UAEM.
- , 1999, *Del Estado gerencial al Estado cívico*, México, UAEM-M.A. Porrúa.
- , 2003, *Gerencia pública en la globalización*, México, UAEM-M.A. Porrúa.
- , 2004, "El mito del nuevo «management» público", en revista *Venezolana de Gerencia*, año 9, núm. 25, enero-marzo Venezuela, Universidad del Zulia.
- HERNÁNDEZ LAOS, E., 1991, *Crecimiento económico y pobreza en México. Una agenda para investigación*, México, UAM.
- HESPEL, Véronique, 2002, "L'avenir institutionnel de l'évaluation des politiques publiques", en *Politiques et Management Public*, vol. 2, núm. 4, número especial, Francia, Institut de Management Public.
- HOFSTÈDE, G., 1980, *Management control of public and not for profit activities*, Austria, International Institute for Applied Systems Analysis.
- HOOD, C., 1994, *Explaining economic policy reversals*, Buckingham, Open University Press.
- y M. Jackson, 1997, *La Argumentación administrativa*, México, FCE.
- INSTITUTO DE DESARROLLO REGIONAL, 2000, *Guía para la evaluación de políticas públicas*, Sevilla, IDR.
- ISAAC, W., 1999, *Dialogue and the art of thinking together, a pioneering approach to communicating in bussines and in life*, Nueva York, Doubleday.
- JOINT COMMITTEE ON STANDARDS FOR EDUCATIONAL EVALUATION, 1994, *The program evaluation standards*, segunda edición, Thousand Oaks, Sage Publications.
- JOBERT, Bruno, 1992, "Représentations sociales. Controverses et débats dans la conduite des politiques publiques", en *Revue Française de Sciences Politiques*, núm. 2, abril.
- , 1998, "La régulation politique, le point de vue d'un politiste", en J. Commailles y B. Jobert (edits.), *Les métamorphoses de la régulation politique*, París, Librairie Général de Droit et de Jurisprudence.
- KESSLER, Marie-Christine et al., 1998, *Evaluation des politiques publiques*, París, L'Harmattan.
- KLIKSBERG, Bernardo, 1989, "¿Cómo formar gerentes sociales? Elementos para el diseño de estrategias", en B. Kliksberg (comp.), *¿Cómo enfrentar la pobreza?*, Argentina, Grupo Editor Latinoamericano.
- , 1994, *Pobreza. El drama cotidiano. Clave para una nueva gerencia social eficiente*, Venezuela, CLAD, PNUD, edit. Tesis-Norma.

- , 1997, *Pobreza. Un tema impostergable*, México, FCE-CLAD-PNUD.
- , 2000a, “Seis tesis no convencionales sobre participación”, en *Acta Sociológica*, núms. 28-29, México, UNAM.
- , 2000b, *Capital social y cultura*, EUA, Instituto de Integración Latinoamericana, INTAL-BID.
- KUHN, Thomas, 1971, *La estructura de las revoluciones científicas*, México, FCE.
- LACASSE, M. François, 1993, “Gestion et émasculación des évaluations, le cas de Canada”, en *Revue Française d'Administration Publique*, num. 66, Francia.
- LAKATOS, Imra, 1989, *La metodología de los programas de investigación científica*, Madrid, Alianza Editorial.
- LASCOURMES, P y M. Setbon, 1996, *L'évaluation pluralista des politiques publiques. Enjeux, pratiques, produits*, Francia, GAPP-CNRS-CGP.
- , 1998, “Pratiques et modèles de l'évaluation” y “Usages et retombées de l'évaluation”, en Marie Christine Kesller *et al.*, *Evaluation des politiques publiques*, París, L'Harmattan.
- LASSWELL, Harold, 1951, “La orientación hacia las políticas”, en Luis F. Aguilar (2000a).
- , 1971, “La concepción emergente de las ciencias políticas”, en Luis F. Aguilar (2000a).
- LAYARD, Richard, 1978, *Análisis costo-beneficio*, México, FCE.
- LECA, Jean *et al.*, 1985, *La théorie politique*, Francia.
- LEEUEW, Frans, Ray Rist y Richard Sonnichsen (eds.), 2000, *Can government learn?*, Colección Comparative Policy Analysis, EUA, Ray Rist editor.
- LINDBLOM, Charles, 1959, “The science of muddling through”, en *Public Administration Review*, vol. 19, EUA.
- , Wood-House, 1993, *The policy making process*, Nueva York, Prentice Hall.
- LÓPEZ PRESA, José Octavio, 2002, “La rendición de cuentas en la política social”, *Serie Cultura de la Rendición de Cuentas*, núm. 3, México, Auditoría Superior de la Federación.
- LYNN, Naomi y Aaron Wildavsky, 1999, *Administración pública. El estado actual de la disciplina*, México, FCE.
- MAJONE, Giandomenico, 1996, “Política pública y administración, ideas, intereses e instituciones”, en Oxford University, *A new handbook of political science*, Oxford, Oxford University Press.
- MANZANOS, César, 1994, “La sociología y su enseñanza. Revisión crítica de algunas cuestiones claves”, en *Memoria docente*, Bilbao, Universidad del País Vasco.

- , 1998, “Introducción a las cuestiones metodológicas en los estudios sociojurídicos”, en *Derecho y Sociedad*, Valencia, Tirant lo Blanch.
- MARTIN, Gérard y Amédine Ruffiot, 2001, “La commande d’ évaluation de politiques sociales territoriales, entre mythes et apprentissages”, en *Revue Politiques et Management Public*, vol. 19, núm. 2, junio, Francia.
- MÉNDEZ, José Luis (comp.), 2000, *Lecturas básicas de administración y políticas públicas*, México, Colmex.
- MEDINA, Alejandro y J. Mejía, 1993, *El control de la implantación de la política pública*, México, Plaza y Valdés.
- MEJÍA LIRA, José, 2003, *La evaluación de la gestión y las políticas públicas*, México, M.A. Porrúa.
- MÉNY, Yves y J.C. Thoenig, 1992, *Las políticas públicas*, Barcelona, edit. Ariel S.A.
- MONNIER, Eric, 1989, “Évaluation des politiques publiques, méthodes et pratiques”, en *Problèmes politiques et sociaux*, núm. 599.
- , 1991, “Objectifs et destinataires des évaluations”, en *Documentación Administrativa*, núm. 224-225, España.
- , 1995, *Evaluación de la acción de los poderes públicos*, España, Instituto de Estudios Fiscales.
- y Vincent Spenlehour, 1992, “L’évaluation dans le triangle de la décision. Opinions singulières et processus pluraliste”, en *Politiques et management publique*, vol. 10, núm. 3, septiembre.
- MULLER, Pierre, 1994, *Les politiques publiques*, Collection Que sais-je?, París, Presses Universitaires de France.
- NACIONES UNIDAS, 2001, *World public sector report, globalization an the state*.
- , 2002, *Informe de Desarrollo Humano*.
- NICKSON, Andrew, 2002, “Transferencia de políticas y reforma en la gestión del sector público. El caso del New Public Management”, en *revista del CLAD, Reforma y Democracia*, núm. 24, Venezuela.
- NIOCHE, Jean Pierre, 1982, “De l’évaluation à l’analyse des politiques publiques”, *Revue Française de Science Politique*, núm. 1, febrero, Francia.
- , 1985, “Les politiques publiques”, en *Traité de Science Politique*, tomo 4, París, PUF.
- , 1993, “L’évaluation de politiques publiques en France, «fast food», recettes du terroir ou cuisine internationale?”, en *Revue Française d’Administration Publique*, núm. 66, abril-junio, Francia.
- y R. Poincard, 1994, *L’évaluation des politiques publiques*, París, Económica.
- NONAKA, I. y T. Nishiguchi, 2001, *Knowledge emergence, social, technical and evolutionary dimensions of knowledge creation*, Nueva York, Oxford University Press.

- OCDE, 1995, *Governance in transition. Public management reforms in OECD countries*.
- , 2001, *Evaluation feedback for effective learning and accountability*.
- O'DONNELL, Guillermo y O. Oszlak, 1976, *Estado y políticas estatales en América Latina, hacia una estrategia de investigación*, Argentina, CEDES.
- OSBORNE Y GABLER, 1994, *La reinvencción del gobierno*, Buenos Aires, Paidós.
- , Plastrik, 1998, *La reducción de la burocracia. Cinco estrategias para reinventar el gobierno*, España, Paidós.
- OSPINA, Sonia, 2001, “Evaluación de la gestión pública, conceptos y aplicaciones en el caso latinoamericano”, en revista *Reforma y Democracia*, núm. 19, Venezuela, CLAD.
- OSUNA, José Luis, 2003, “La evaluación participativa, una visión integral y pluralista”, en Bañón (comp.), *La evaluación de la acción y de las políticas públicas*, Madrid, Díaz de Santos.
- OSZLAK, Óscar, 1980, *Políticas públicas y regímenes políticos, reflexiones a partir de algunas experiencias latinoamericanas*, Estudio CEDES, vol. 3, núm. 2, Argentina.
- , 2003, “Profesionalización de la función pública en el marco de la Nueva Gestión Pública”, en David Arellano, Rodrigo Egaña, Óscar Oszlak y Regina Pacheco, *Retos de la profesionalización de la función pública*, Venezuela, CLAD-AECI-MAP-FIIAPP.
- PAPADOPOULOS, Yannis, 2001, “Transformations du style de l'action publique et responsabilité politique”, en *Politiques et Management Public*, vol. 19, núm. 1, marzo, Francia, Institut de Management Public.
- PARAMÉS, Carlos, 1978, *Introducción al management*, España, ENAP.
- PARDO, María del Carmen, 2000, “La evaluación gubernamental, prioridad política de 1982 a 1988”, en José Luis Méndez (comp.), *Lecturas básicas de administración y políticas públicas*, México, Colmex.
- PATTON, M.Q., 1997, *Utilization-focused evaluation. The new century text*, Thousand Oaks, Sage Publications.
- PAULA, L.A. Maciel de y J. Cléber, 2002, “La utilización de indicadores sociales en la evaluación de políticas públicas. Estudio de un caso de reforma agraria”, en revista *Reforma y Democracia*, núm. 24, Venezuela, CLAD.
- PERRET, Bernard, 1993, “Refléxions sur les différent modèles d'évaluation”, en *Revue Française d'Administration Publique*, núm. 66, Francia.
- POLLIT, Christopher, 1993, *Managerialism and the public services*, Oxford, Blakwell.
- , 1998, “Institutions et usages de l'évaluation au Royaume Uni, une vue d'ensemble”, en Conseil Scientifique de l'Évaluation, *L'évaluation en développement 1997*, Francia, La Documentation Française.

- POPPER, K., 1977, *La lógica de la investigación científica*, Madrid, Tecnos.
- QUADE, E., 1975, *Analysis for public decisions*, EUA, American Elsevier.
- RIEPER Y TOULEMONDE, 1996, *Politics and practices of inter-governmental evaluation*, Comparative Policy Analysis, Series Editor.
- RIST, Ray, 1990, *Program evaluation and the management of government, patterns and prospects across eight nations*, New Brunswick, Transaction Publishers.
- ROSSI y Freeman, 1993, *Evaluation: a systematic approach*, Sage Publications.
- RUANO, José Manuel, 2003, “La doble naturaleza técnica y política de la evaluación del rendimiento”, en Bañón, 2003, *La Evaluación de la acción y de las políticas públicas*, Madrid, Díaz de Santos.
- RUBIO, M. José y Jesús Varas, 1999, *El análisis de la realidad en la intervención social*, Madrid, Editorial CCS.
- SAINT-CLAIR, W., 1993, “L'évaluation des politiques publiques au Royaume Uni”, en *Revue Française d'Administration Publique*, núm. 66, abril-junio, Francia.
- SANÍN ÁNGEL, Héctor, 1999, “Control de gestión y evaluación de resultados en la gerencia pública”, *Serie Manuales*, núm. 3, agosto, Santiago de Chile, CEPAL.
- SCOTT, R., 1995, *Institutions and organizations*, Thousand Oaks, Sage, Publications.
- SCHEIN, E.H., 1985, *Organizational Psychology*, Englewood Cliffs, Nueva Jersey, Prentice Hall.
- SCHWANDT, Thomas y Edward Halpern, 1988, *Linking auditing and metaevaluation*, Newbury Park, Sage Publications.
- SCRIVEN, Michael, 1967, “The methodology of evaluation”, en Tyler *et al.*, *Perspective of curriculum evaluation*, Chicago, Rand McNally.
- , 1969, “An introduction to meta-evaluation”, en *Educational Product Report*.
- , 1973, “Goal free evaluation”, en E. House (ed.), *The politics and process*, Nueva York, McCutchan.
- , 1976, “Evaluation bias and its control”, en *Evaluation Studies Review Annual*, vol. 1.
- , 1991, *Evaluation Thesaurus*, Newbury Park, Sage Publications.
- SEN, Amartya, 1999, *Development as freedom*, Nueva York, Knopf Press.
- SETBON, Michel, 1998, “Introduction”, en Marie Christine Kesller *et al.*, *Evaluation des politiques publiques*, París, L'Harmattan.
- SIMON, Herbert, 1972, *El comportamiento administrativo*, España, Aguilar, 1ra edición, 1947.
- , 1957, *Models of man, social and rational*, Nueva York, Wiley.

- SKOUFIAS, Emmanuel, 2001, *Progreso and its impacts on the human capital and welfare of households in rural Mexico; a synthesis of the results of an evaluation by IFPRI*, Washington, D.C., International Food Policy Research Institute.
- y Susan Parker, 2001, “Conditional cash transfers and the impact of child work and schooling, evidence from the Progreso program in Mexico”, en *Economía: Journal of Latin America and the Caribbean Economic Association*, vol. 2, núm. 1.
- SOLARTE PAZOS, Leonardo, 2004, *La evaluación de políticas públicas en el Estado liberal*, Cali, Colombia, Universidad del Valle.
- SOLÍS, Leopoldo, 1987, *La realidad económica mexicana, retrovisión y perspectivas*, México, Siglo XXI.
- STAME, Nicoletta, 1998, “L'évaluation en Italie”, en Conseil Scientifique de l'Évaluation, *L'évaluation en développement 1997*, Francia, La Documentation Française.
- STARLING, Grover, 1979, *The politics and economics of public policy*, EUA, The Dorsey Press.
- STEINBRUNER, John, 1974, *The cybernetic theory of decision*, Nueva Jersey, Princeton University Press.
- STOCKEY Y ZECKHAUSER, 1978, *A primer for policy analysis*, EUA, W.W. Norton and Co.
- STUFFLEBEAM, Daniel y Anthony Shinkfield, 1987, *Evaluación sistemática*, España, Paidós.
- SUBIRATS, Joan, 1989, *Análisis de políticas públicas y eficacia de la administración*, España, INAP.
- SULBRANDT, José, 1993, “La evaluación de los programas sociales, una perspectiva crítica de los modelos actuales”, en Bernardo Kliksberg (comp.), *Pobreza. Un tema impostergable. Nuevas propuestas a nivel mundial*, México, CLAD-FCE-PNUD.
- TAYLER, R., 1967, “Changing concepts of educational evaluation”, en R.E. Stake (comp.), *Perspectives of curriculum evaluation*, vol. 1, Nueva York, Rand McNally.
- TARRÉS, María Luisa, 2001, *Observar, escuchar y comprender. Sobre la tradición cualitativa de la investigación social*, México, Colmex-Flacso-M.A. Porrúa.
- THOENIG, Jean-Claude, 1994, “La gestion systémique de la sécurité publique”, en *Revue Française de Sociologie*, XXXV, Francia.
- , 1998a, “Que peut-on apprendre s'agissant de la France de la rencontre franco-britannique?”, en Conseil Scientifique de l'Évaluation, *L'évaluation en développement 1997*, Francia, La Documentation Française.



- , 1998b, “L'évaluation de l'action publique locale”, en Marie Christine Kesller *et al.*, *Evaluation des politiques publiques*, París, L'Harmattan.
- , 2002, “La evaluación en actos. Lecciones del pasado y preguntas para el futuro”, en revista *Gestión y Política Pública*, vol. XI, núm. 2, segundo semestre, México, CIDE; y, 2002, versión en francés, *Politiques et Management Public*, vol. 2, núm. 4, número especial, diciembre, Francia, Institut de Management Public.
- TORRES, Lourdes y Vicente Pina, 2003, “Construcción de indicadores para la evaluación de políticas públicas”, en Bañón (comp.), *La evaluación de la acción y de las políticas públicas*, Madrid, Díaz de Santos.
- TORRES SALCIDO, Gerardo, 2003, “Introducción: las políticas de abasto popular de alimentos. Estado actual y perspectivas”, en Gerardo Torres Salcido (coord.), *Políticas de abasto alimentario. Alternativas para el Distrito Federal y su zona metropolitana*, México, ediciones Casa Juan Pablos, Facultad de Ciencias Políticas y Sociales, UNAM.
- TURGEON, Jean, 1998, “Le programme de l'action communautaire pour les enfants”, en Marie Christine Kesller *et al.*, *Evaluation des politiques publiques*, París, L'Harmattan.
- TROSA, Sylvie, 1993, “Un premier bilan de l'expérience française”, en *Revue française D'administration publique*, núm. 66, abril-junio, Francia.
- UAM-A y Fundación Friedrich Ebert, 1997, “Mujer y políticas públicas”, revista *El Cotidiano*, núm. 84, julio-agosto, México.
- UAM-X, de 1996, “Políticas públicas y sociedad”, *revista de Política y Cultura*, núm. 7, otoño, México.
- UVALLE, Ricardo, 1994, “Prólogo”, en Adriana Plasencia, *Gerencia pública. Bases teórico-conceptuales*, México, Colegio Nacional de Ciencias Políticas y Administración Públicas, A.C.
- , 1997, “Estudio introductorio”, en C. Hood, y M. Jackson, *La argumentación administrativa*, México, FCE.
- , 1998, “El carácter multifacético de la gestión pública contemporánea”, en *revista del Instituto de Administración Pública del Estado de México*, núm. 37, enero-marzo. Toluca, México, INAPEM.
- , 2001, “El fortalecimiento de la vida democrática, la nueva relación de la sociedad y el Estado”, en revista *Reforma y Democracia*, núm. 20, junio, Caracas, Venezuela, CLAD.
- VEDUNG, E., 1997, *Evaluación de políticas públicas y programas*, Madrid, Ministerio de Trabajo y Asuntos Sociales, Instituto Nacional de Servicios Sociales.

- WEISBROD, 1978, en Richard Layard (comp.), *Análisis costo-beneficio*, México, FCE.
- WEISMER, Eduardo, 2000, “Función de evaluación de planes, programas, estrategias y proyectos”, en *Series CEPAL*, núm. 4, mayo. Santiago de Chile, CEPAL.
- WEISS, Carol, 1992, *Investigación evaluativa*, México, Trillas.
- , 1998, *Evaluation. Methods for studying programs and policies*, EUA, Prentice Hall.
- , 1999, “Las investigaciones de políticas, ¿datos, ideas y argumentos?”, en Peter Wagner, Carol Weiss, Bjorn Wittrock y Hellmut Wollman (comp.), *Ciencias sociales y estados modernos. Experiencias nacionales e incidencias teóricas*, México, FCE.
- WILSON, Woodrow, 1887, “The study of administration”, en *Political Science Quarterly*, vol. II, núm. 2, junio.
- ZALTSMAN, Ariel, 2004, “La evaluación de resultados en el sector público argentino, un análisis a la luz de otras experiencias en América Latina”, revista del CLAD, *Reforma y Democracia*, núm. 29, junio.
- ZICCARDI, Alicia y Homero Saltalamachia, 1997, *Metodología de evaluación del desempeño de los gobiernos locales en ciudades mexicanas*, México, Instituto de Investigaciones Sociales, UNAM.

Documentos, ponencias, periódicos,  
tesis y consultas a Internet

- ACUERDO PARLAMENTARIO PARA LA ORGANIZACIÓN Y REUNIONES DE LAS COMISIONES Y COMITÉS DE LA CÁMARA DE DIPUTADOS, [www.diputados.gob.mx](http://www.diputados.gob.mx), consulta efectuada el 2 de mayo de 2003.
- ALAMILLA CRUZ, Marina, s/f, “Evaluación externa de los programas federales con reglas de operación”, *Mimeo*.
- ARELLANO, David y José Ramírez, “Presupuesto dirigido a resultados, los dilemas de la evaluación por desempeño en contextos latinoamericanos. Lecciones desde México”, en [www.clad.org.ve](http://www.clad.org.ve), consulta efectuada el 26 de julio de 2003.
- AUDITORÍA SUPERIOR DE LA FEDERACIÓN, “Informe de resultados sobre la revisión de la cuenta de la hacienda pública federal 2000”, en [www.asf.gob.mx](http://www.asf.gob.mx), consulta efectuada el 13 de septiembre de 2004.
- , “Informe del resultado de la revisión y fiscalización superior de la cuenta pública 2001”, en [www.asf.gob.mx](http://www.asf.gob.mx), consulta efectuada el 13 de septiembre de 2004.

- , “Informe del resultado de la revisión y fiscalización superior de la cuenta pública 2002”, en [www.asf.gob.mx](http://www.asf.gob.mx), consulta efectuada el 13 de septiembre de 2004.
- BAKER, Judy, 2000, *Evaluación del impacto de los proyectos de desarrollo en la pobreza. Manual para profesionales*, Banco Mundial, [www.worldbank.org](http://www.worldbank.org), consulta realizada el 10 de mayo del 2003.
- BANCO MUNDIAL, 2004b, “México presenta avances y enfrenta retos en su lucha contra la pobreza”, nota publicada el 28 de julio, en [www.worldbank.org](http://www.worldbank.org), consulta efectuada el 29 de julio de 2004.
- BARRERA CHAVIRA, Edith, “La evaluación del ejecutivo por el legislativo, el caso de los programas compensatorios del sector educativo”, [www.eddheu.gob.mx](http://www.eddheu.gob.mx), consulta efectuada el 10. de mayo de 2003.
- BID, 1997, “Evaluación, una herramienta de gestión para mejorar el desempeño de los proyectos”, Washington, *Oficina de Evaluación del BID*.
- BOLTVINIK, Julio, 1998, “¡15 millones de pobres extremos!”, en periódico *La Jornada*, 16 de octubre.
- , 2002, “Hacia dónde va la Sedeso”, en periódico *La Jornada*, 5 de abril.
- , 2003, “Evaluaciones *light* del Oportunidades”, en periódico *La Jornada*, 25 de abril.
- , 2005, “Sedeso boicotea la Ley de Desarrollo Social”, en periódico *La Jornada*, 26 de agosto.
- , 2005, “Evaluación, dos modelos”, en periódico *La Jornada*, 2 de septiembre.
- , 2005, “Voluntad divina en bueyes ajenos”, en periódico *La Jornada*, 23 de septiembre.
- , 2005, “¿Controversia constitucional o acuerdo?”, en periódico *La Jornada*, 30 de septiembre.
- , 2005, “Acuerdos entre Sedeso y el Legislativo”, en periódico *La Jornada*, 7 de octubre.
- , 2005, “Reglamento de la Ley de Desarrollo Social”, en periódico *La Jornada*, 14 de octubre.
- , 2005, “Desastre en el Oportunidades”, en periódico *La Jornada*, 4 de noviembre.
- , 2006, “Sedeso, crisis e ilegalidad”, en periódico *La Jornada*, 20 de enero.
- , 2006, “El reto de la evaluación en México”, en periódico *La Jornada*, 10 de febrero.
- , 2006, “Elecciones excluyentes y controversia constitucional”, en periódico *La Jornada*, 17 de febrero.

- , 2006, “Imposición del modelo social único”, 24 de febrero.
- , 2006, “Fortalecer lo social frente a lo económico”, en periódico *La Jornada*, 3 de marzo.
- , 2006, “La Sedeso quiere medir la pobreza”, en periódico *La Jornada*, 10 de marzo.
- BUSTELO, María, 2001, *La evaluación de las políticas públicas de igualdad de género de los gobiernos central y autonómicos en España, 1995-1999*, tesis doctoral, Madrid, Universidad Complutense.
- , 2002, “Meta-evaluation as a tool for the empowerment and development of the evaluation function in Public Administration”, ponencia presentada en la *V Conferencia Europea de Evaluación, Sevilla*, [www.europeanevaluation.org](http://www.europeanevaluation.org).
- CARDOZO BRUM, Myriam, 5 al 9 de noviembre de 2001, “Resistencias y problemas para la evaluación de la política social mexicana”, en *VI Congreso Internacional del CLAD sobre la Reforma del Estado y la Administración Pública*, Buenos Aires, Argentina.
- , 28 al 31 de octubre de 2003, “Evaluación de las políticas sociales en México”, *VIII Congreso Internacional del CLAD sobre la Reforma del Estado y la Administración Pública*, Panamá, Panamá.
- , 2 al 5 de noviembre de 2004, “Metaevaluación de programas sociales desarrollados por el gobierno mexicano”, *IX Congreso Internacional del CLAD sobre la Reforma del Estado y la Administración Pública*, Madrid, España.
- CENTENO, Helen, 2002, *Meta-evaluación del Programa Antidrogas y Salud Sexual y Reproductiva para Adolescentes, Jamaica 1999-2002, lecciones aprendidas y contribución metodológica*, tesis del Master Profesional en Población y Salud, Costa Rica, Universidad de Costa Rica.
- CENTRE FOR EUROPEAN EVALUATION EXPERTISE, 2002, [www.eureval\\_c3e.fr](http://www.eureval_c3e.fr), página consultada el 4 de noviembre.
- CERTU, 2002, “Évaluation des politiques publiques, faut-il quantifier pour évaluer?”, cuaderno núm. 6, Francia, [www1.certu.fr](http://www1.certu.fr), consulta efectuada el 4 de noviembre
- CIDE, “Evaluación de los impactos de Progresía sobre los hogares, Encuestas de Evaluación, (ENCCEL)”, en [www.progresia.gob.mx](http://www.progresia.gob.mx), consulta realizada el 2 de mayo de 2003.
- , “Evaluación del Programa de Educación, Salud y Alimentación, Progresía) a partir de: Indicadores de Seguimiento, Evaluación y Gestión 1998-2001 y encuestas de Evaluación 2000”, en [www.progresia.gob.mx](http://www.progresia.gob.mx), consulta realizada el 27 de agosto de 2004.

- CIESAS-Istmo, 2003, “Resultados de la evaluación de gabinete y campo del programa de Planeación, Capacitación y Organización Social, enero-diciembre de 2002”, en *CD entregado por el IFAI*.
- , 2003, “Evaluación final de gabinete y campo del programa de operación de Albergues Escolares Indígenas de enero-diciembre de 2002”, en *CD entregado por el IFAI*.
- , 2003, “Evaluación final de gabinete y campo del programa de Bienestar Social”, en *CD entregado por el IFAI*.
- , 2003, “Evaluación final del programa de Reparación y Equipamiento de Albergues Escolares del INI, enero-diciembre de 2002”, en *CD entregado por el IFAI*.
- , Occidente, 2002, “Documento final de diagnóstico cualitativo de hogares semiurbanos. Localidades de 2,500 a 50,000 habitantes, 2001”, w [www.oportunidades.gob.mx](http://www.oportunidades.gob.mx), consulta efectuada el 8 de octubre de 2003.
- , Occidente, 2003, “Evaluación cualitativa del programa Desarrollo Humano Oportunidades. Seguimiento de impacto 2001-2002”, en *CD entregado por el IFAI*.
- , Occidente e INSP, 2003, “Evaluación de resultados de impacto del programa Desarrollo Humano Oportunidades”, en *CD entregado por el IFAI*.
- COHEN, Ernesto, 2004, “Evaluación y monitoreo de programas sociales”, seminario núm. 22, octubre, Sedesol.
- COLEGIO DE MÉXICO, 2003, “Evaluación externa del programa para la Superación de la Pobreza Urbana. Reporte final”, en *CD entregado por el IFAI*.
- COLEGIO NACIONAL DE CIENCIAS POLÍTICAS Y ADMINISTRACIÓN PÚBLICA, A.C., 2003, “Investigación para el Desarrollo Local. Informe final y anexos del periodo enero-diciembre del 2002”, en *CD entregado por el IFAI*.
- COMIM, Flavio, 2004, “Introducción a las capacidades básicas planteadas por Amartya Sen y su aplicación en el contexto mexicano”, conferencia impartida el 29 de marzo en el *Congreso de Desarrollo Social y Libertad sobre Pobreza y Capital Social 2004*, Sedesol-UIA.
- CONSEJO CIENTÍFICO DEL CLAD, octubre de 1998, “Una nueva gestión pública para América Latina”, [www.clad.org.ve](http://www.clad.org.ve), consulta efectuada el 4 de octubre de 2004.
- CONSTITUCIÓN POLÍTICA DE LOS ESTADOS UNIDOS MEXICANOS, en [www.cddhuc.gob.mx](http://www.cddhuc.gob.mx), consulta efectuada el 22 de marzo de 2004.
- CUNILL, Nuria, 2000, “Responsabilización por el control social”, en Consejo Científico del CLAD (coord.), *La responsabilización en la nueva gestión pública*

*latinoamericana*, Buenos Aires, CLAD-BID-Editorial Universitaria de Buenos Aires.

EES, “Tres movimientos en la evaluación contemporánea, aprendizaje, teoría y pruebas”, convocatoria a la V Conferencia Bienal, España, octubre 2002, [www.europeanevaluation.org](http://www.europeanevaluation.org)

EL COLEGIO MEXIQUENSE, 2003a, “Coinversión Social 2001-2002. Informe de resultados definitivos”, en *CD entregado por el IFAI*.

—————, 2003b, “Informe de resultados definitivos de la evaluación correspondiente al periodo enero-diciembre de 2002 del Programa de Desarrollo Económico Productivo de los Pueblos Indígenas a cargo del INI”, en *CD entregado por el IFAI*.

ENGEL, Paul y Charlotte Carlsson, 2002, “Enhancing learning through evaluation, approaches, dilemmas and some posible ways forward”, ponencia presentada en la *V Conferencia Europea de Evaluación, Sevilla*, [www.europeanevaluation.org](http://www.europeanevaluation.org).

ESCUELA NACIONAL DE ANTROPOLOGÍA E HISTORIA, INAH, 2003, “Evaluación externa del programa Promoción y Procuración de Justicia del Instituto Nacional Indigenista. Informe Final”, en *CD entregado por el IFAI*.

ESCUELA NACIONAL DE TRABAJO SOCIAL, UNAM, 2003, “Evaluación externa. Programa de Capacitación para el Fortalecimiento de la Sociedad Civil y la Participación Social. Informe final”, en *CD entregado por el IFAI*.

—————, “Evaluación del programa de Atención a Jornaleros Agrícolas. Cuarto Informe”, en *CD entregado por el IFAI*.

FACULTAD DE ECONOMÍA, U.A. DE NUEVO LEÓN, 2003, “Evaluación de los programas de la Conaza en el año 2002”, en *CD entregado por el IFAI*.

FACULTAD DE ECONOMÍA, UNAM, 2003, “Evaluación externa del programa de Ahorro y Subsidios para la Vivienda Progresiva”, en *CD entregado por el IFAI*.

FACULTAD DE ECONOMÍA, UNAM, 2003, “Evaluación externa 2002. Programa Oportunidades Productivas”, en *CD entregado por el IFAI*.

FORESTI, Marta *et al.*, 2002, “The role of civil society in evaluating poverty reduction strategies and producing evidence for policy development, myth or reality?” ponencia presentada en la *V Conferencia Europea de Evaluación, Sevilla*, [www.europeanevaluation.org](http://www.europeanevaluation.org)

FRANCO, Rolando, 28-31 de octubre de 2003, “Descentralización, participación y competencia en la gestión social”, ponencia presentada en el *VIII Congreso Internacional del CLAD, Panamá, Panamá*.

FUNSALUD, 2003, “Resultados del programa Expertos en Acción”, en *CD entregado por el IFAI*.

- GARCÍA SÁNCHEZ, Ester, 2003, “La evaluación de programas de reforma educativa en España. Tres estudios de caso desde un enfoque de metaevaluación”, tesis doctoral, Universidad Complutense de Madrid.
- , “Metaevaluación”, en *Diccionario crítico de ciencias sociales*, España, www.ucm.es.
- , 2004, “La evaluación educativa en España, balance y metaevaluación”, ponencia presentada al *IX Congreso Internacional del CLAD*, Madrid, 2 al 5 de noviembre.
- GEA, 2003, “Evaluación externa del programa de Abasto Rural operado por Diconsa”, www.diconsa.gob.mx, consulta efectuada el 8 de octubre de 2003.
- GUERRERO, Melchor, María Antonia Cobacho y Encarnación Martínez, 2001, “Metodología para la evaluación de logros de desarrollo rural”, en *Iniciativa Comunitaria LEADER II*, España, www.eclac.cl, consulta realizada el 2 de mayo de 2003.
- GUINART I SOLÁ, Josep María, 28-31 de octubre de 2003, “Indicadores de gestión para las entidades públicas”, ponencia presentada al *VIII Congreso Internacional del CLAD*, Panamá, Panamá.
- HERNÁNDEZ LICONA, Gonzalo, mayo de 2003, “Sistema de evaluación y monitoreo”, *trabajo presentado en la Conferencia Internacional Mejores Prácticas de Política Social*, México, D.F., Sedesol-Banco Mundial-BID-CEPAL.
- , 2005, “The political side of social program evaluation”, trabajo inédito, México, Sedesol.
- HINTZE, Jorge, 28-31 de octubre de 2003, “Responsabilización, instrumentos de rendición de cuentas y evaluación de gestión del valor público”, ponencia presentada al *VIII Congreso Internacional del CLAD*, Panamá, Panamá.
- IFPRI, “Mexico, Evaluation of Progres, 1997-1999”, en www.ifpri.org, consulta efectuada el 15 de mayo de 2003.
- INDETEC, 2003, “Evaluación del programa de Fortalecimiento Institucional para la Gestión Social Municipal 2002”, en *CD entregado por el IFAI*.
- INEGI, 2000, *Niveles de bienestar en México*, en www.inegi.gob.mx, fecha de consulta, 23 de octubre de 2002.
- INSP-CIESAS, 2003, “Resumen ejecutivo de la evaluación externa del Programa de Desarrollo Humano Oportunidades 2002”, www.oportunidades.gob.mx, consulta efectuada el 8 de octubre de 2003.
- INSP-CIDE, 2003, “Evaluación del impacto de Oportunidades sobre la inscripción escolar, primaria, secundaria y media superior”, en *CD entregado por el IFAI*.

- , 2003, “Análisis sobre el cambio en variables demográficas y económicas de los hogares beneficiarios del programa oportunidades 1997-2002”, en *CD entregado por el IFAI*.
- INSP, 2003, “Evaluación del efecto de Oportunidades sobre la utilización de servicios de salud en el medio rural”, en *CD entregado por el IFAI*.
- , 2003, “Evaluación de la identificación de familias beneficiarias en el medio urbano”, en *CD entregado por el IFAI*.
- , 2003, “Evaluación del cumplimiento de metas, costos unitarios y apego del programa Oportunidades a las Reglas de Operación”, en *CD entregado por el IFAI*.
- INSTITUTO MAYA, 2003a, “Informe final de la evaluación externa de los Programas Regionales para Zonas de Alta Marginación e Indígenas”, en *CD entregado por el IFAI*.
- , 2003b, “Informe final de la evaluación externa del Programa Microrregiones”, en *CD entregado por el IFAI*.
- ITESM, 2003a, “Evaluación de resultados del Programa de Abasto Social de Leche, a cargo de Liconsa, S.A. de C.V. Resumen ejecutivo al reporte final de resultados de la evaluación enero a diciembre de 2002”, [www.liconsa.gob.mx](http://www.liconsa.gob.mx), consulta efectuada el 8 de octubre de 2003.
- , 2003b, “Evaluación de resultados del Programa Tortilla, a cargo de Liconsa, S.A. de C.V. Resumen ejecutivo de la evaluación final enero-diciembre de 2002”, [www.liconsa.gob.mx](http://www.liconsa.gob.mx), consulta efectuada el 8 de octubre de 2003.
- JASO, Sofía y Mauricio Merino, diciembre de 1994, “Solidaridad, los matices de una política”, documento preparado para el Comité Técnico de Evaluación del Programa Nacional de Solidaridad, mimeo.
- KLIKSBERG, Bernardo, 28-31 de octubre de 2003, “Nuevas ideas en política social y concertación para el desarrollo social”, ponencia presentada en el *VIII Congreso Internacional del CLAD*, Panamá, Panamá.
- LAJODIE, B. y J. Toulemonde, junio de 2002, “Méta-évaluation, bilan des évaluations réalisées en Limousin au cours de la période 1994-1998”, ponencia presentada a las *Cuartas Jornadas de la Société Française d'Évaluation*, Lille, France.
- LARRÚ, José María, 2003, “Análisis de los resultados y metaevaluación del sistema de evaluaciones de EuropeAid”, España, CECOD, [www.cecod.net](http://www.cecod.net), consulta efectuada el 22 de junio de 2004.
- LEY DE ADQUISICIONES, ARRENDAMIENTOS Y SERVICIOS DEL SECTOR PÚBLICO, *Diario Oficial de la Federación* del 4 de enero de 2000.



- LEY DE FISCALIZACIÓN SUPERIOR DE LA FEDERACIÓN, [www.asf.gob.mx](http://www.asf.gob.mx), consulta efectuada el 2 de mayo de 2003.
- LEY FEDERAL DE ACCESO A LA INFORMACIÓN PÚBLICA GUBERNAMENTAL, [www.secodam.gob.mx](http://www.secodam.gob.mx), consulta efectuada el 2 de mayo de 2003.
- LEY FEDERAL DE FOMENTO A LAS ACTIVIDADES REALIZADAS POR LAS ORGANIZACIONES DE LA SOCIEDAD CIVIL, [www.cddhcu.gob.mx](http://www.cddhcu.gob.mx), consulta efectuada el 2 de mayo de 2003.
- LEY GENERAL DE DESARROLLO SOCIAL, [www.sedesol.gob.mx](http://www.sedesol.gob.mx), consulta efectuada el 31 de marzo de 2004.
- LEY ORGÁNICA DEL CONGRESO GENERAL DE LOS ESTADOS UNIDOS MEXICANOS, [www.camaradediputados.gob.mx](http://www.camaradediputados.gob.mx), consulta efectuada el 2 de mayo de 2003.
- LICHA, Isabel, octubre de 1998, "Participación comunitaria", mimeo., EUA, INDES-BID.
- , mayo de 2001, "Participación ciudadana y gerencia social", mimeo., EUA, INDES-BID.
- LONGO, Francisco y Koldo Echevarría, 2001, "La nueva gestión pública en la reforma del núcleo estratégico del gobierno. Experiencias latinoamericanas", en CLAD, Serie Estudios de Casos, *Nueva Gestión Pública y regulación en América latina. Balances y desafíos*, [www.clad.org.ve](http://www.clad.org.ve), consulta realizada el 4 de octubre de 2004.
- LÓPEZ CALVA, Luis F 2003, "Medición multidimensional de la pobreza en México", seminario impartido en la *Serie sobre Desarrollo Social*, Subsecretaría de Prospectiva, Planeación y Evaluación, Sedesol, el 10. de octubre de 2003.
- MOLINET, Jonathan, 28-31 de octubre de 2003, "¿Las reglas de operación en programas de política social producen efectos inesperados?", ponencia presentada en el *VIII Congreso Internacional del CLAD*, Panamá, Panamá.
- MONNIER, Eric y Pitarelli, 2000, "Benchmarking; the missing link between evaluation and management", ponencia presentada en la *IV Conferencia Europea de Evaluación*, Lausanne, [www.europeanevaluation.org](http://www.europeanevaluation.org).
- MUÑOZ, Óscar T., "La evaluación de la actuación del Poder Ejecutivo de la Federación por las Comisiones de Dictamen Legislativo de la H. Cámara de Diputados", en [www.eddheu.gob.mx](http://www.eddheu.gob.mx), consulta realizada el 10. de mayo de 2003.
- NIOCHE, Jean Pierre, junio de 1992, Conferencia sobre evaluación, impartida en CIDE A.C., México.
- NIRENBERG, Olga, 28-31 de octubre de 2003, "El rol del Estado para la participación social en la evaluación", ponencia presentada en el *VIII Congreso Internacional del CLAD*, Panamá, Panamá.

- OLMEDO, Raúl, 28-31 de octubre de 2003, “Sociedad y Estado en la complejidad creciente de los asuntos públicos”, ponencia presentada en el *VIII Congreso Internacional del CLAD*, Panamá, Panamá.
- ONU, marzo de 1994, “Informe de la undécima reunión de expertos sobre el Programa de las Naciones Unidas en materia de administración y finanzas”, Nueva York, ONU.
- PETERS, Guy, 1998, “Globalization, Institutions and governance”, Paper 51, Robert Schuman Centre, Instituto Univesitario Europeo de Florencia.
- PITARELLI, E. y E. Monnier, 2000, “Benchmarking: the missing link between evaluation and Management?”, ponencia presentada en la *IV Conferencia europea de evaluación*, Lausanne, [www.europeanevaluation.org](http://www.europeanevaluation.org)
- PNUD, 2004, “Democracia en América Latina, hacia una democracia de ciudadanos y ciudadanas”, en [www.undp.org](http://www.undp.org), consulta efectuada el 10 de octubre de 2004.
- PLAN NACIONAL DE DESARROLLO 2001-2006, en [www.sedesol.gob.mx](http://www.sedesol.gob.mx), consulta efectuada el 10. de mayo de 2003.
- POLITIQUES PUBLIQUES URBAINES, bibliographie, en [www.urbanisme.equipement.gouv.fr](http://www.urbanisme.equipement.gouv.fr), consulta efectuada el 10. de mayo de 2003.
- PRESUPUESTO DE EGRESOS DE LA FEDERACIÓN PARA EL EJERCICIO FISCAL DEL AÑO 2000, publicado en el *Diario Oficial de la Federación*, primera sección del 31 de diciembre de 1999.
- PRESUPUESTO DE EGRESOS DE LA FEDERACIÓN PARA EL EJERCICIO FISCAL DEL AÑO 2001, publicado en el *Diario Oficial de la Federación*, primera sección del 31 de diciembre de 2000.
- PRESUPUESTO DE EGRESOS DE LA FEDERACIÓN PARA EL EJERCICIO FISCAL DEL AÑO 2002, publicado en el *Diario Oficial de la Federación*, primera sección, 10. de enero de 2002.
- PROGRAMA INSTITUCIONAL OPORTUNIDADES 2001-2006, en [www.sedesol.gob.mx](http://www.sedesol.gob.mx), consulta efectuada el 10. de mayo de 2003.
- PROGRAMA NACIONAL DE DESARROLLO SOCIAL 2001-2006, en [www.sedesol.gob.mx](http://www.sedesol.gob.mx), consulta efectuada el 10. de mayo de 2003.
- PROGRESA, [www.progresa.gob.mx](http://www.progresa.gob.mx), consulta realizada el 2 de mayo de 2003.
- PROGRAMA PARA EL DESARROLLO ECONÓMICO PRODUCTIVO DE LOS PUEBLOS Y COMUNIDADES INDÍGENAS 2002, Evaluación externa, en *CD entregado por el IFAI*.
- QUERMONE Y ROUBAN, 1983, “L'évaluation des politiques publiques et le contrôle politique”, en *Coloquio L'évaluation des politiques publiques*, París.

- REGLAMENTO PARA EL GOBIERNO INTERIOR DEL CONGRESO GENERAL DE LOS ESTADOS UNIDOS MEXICANOS, [www.diputados.org.mx](http://www.diputados.org.mx), consulta realizada el 22 de junio de 2003.
- REGLAMENTO INTERIOR DE LA UNIDAD DE EVALUACIÓN Y CONTROL DE LA COMISIÓN DE VIGILANCIA DE LA CÁMARA DE DIPUTADOS, [www.shcp.gob.mx](http://www.shcp.gob.mx), consulta efectuada el 1o. de mayo de 2003.
- REPETTO, Fabián, 28-31 de octubre de 2003, “Capacidad estatal, requisito necesario para una mejor política social en América Latina”, ponencia presentada en el *VIII Congreso Internacional del CLAD*, Panamá, Panamá.
- REYNA, José Luis y Alberto García Rocha (coords.), s/f, “Metodología para la evaluación del impacto socioeconómico del Programa Nacional de Solidaridad”, mimeo., Colmex.
- ROSENBAUM, Allan, 2002, “Program evaluation and inequality, on the need for social equity impact análisis”, *VII Congreso Internacional del CLAD sobre la Reforma del Estado y la Administración Pública*, Lisboa, 8-11 de octubre.
- SEDESOL, [www.sedesol.gob.mx](http://www.sedesol.gob.mx), consulta efectuada el 2 de mayo de 2003.
- , julio de 2002, “Medición de la Pobreza”, [www.sedesol.gob.mx](http://www.sedesol.gob.mx), el 2 de mayo de 2003.
- , *Reglas de Operación, Bases y Términos de Referencia de diversos programas sociales*, mimeo.
- , marzo de 2004, “Reconoce el Banco Mundial a Sedesol como líder en evaluación de programas sociales”, comunicado núm. 048/2004, en [www.sedesol.gob.mx](http://www.sedesol.gob.mx), consulta efectuada el 4 de marzo de 2004.
- , 25 de agosto de 2005, “Decreto que regula el Consejo Nacional de Evaluación de la Política de Desarrollo Social”, *Diario Oficial de la Federación*.
- , 18 de enero de 2006, “Reglamento de la Ley General de Desarrollo Social”, *Diario Oficial de la Federación*.
- SFP, julio de 2003, “Guía para realizar evaluaciones externas según las disposiciones emitidas por la SECODAM y la SHCP y publicadas en el *Diario Oficial de la Federación* el 30 de abril de 2002”, mimeo.
- SHCP-SECODAM, 30 de abril de 2002, “Acuerdo por el que se expiden los requisitos mínimos que las dependencias y entidades de la Administración Pública Federal deberán considerar para seleccionar a las instituciones académicas y de investigación u organismos especializados, de carácter nacional o internacional, interesados en realizar la evaluación de resultados de los programas gubernamentales sujetos a Reglas de Operación en el ejercicio fiscal 2002”, *Diario Oficial de la Federación*, segunda sección, pp. 4-7.

- SHCP, 2003, "Superación de la pobreza", en Cuenta de la Hacienda Pública 2002, [www.shcp.gob.mx](http://www.shcp.gob.mx), consulta efectuada el 22 de octubre de 2003.
- SKOUFIAS, Emmanuel, "Importancia de la evaluación de los programas sociales, el ejemplo de Oportunidades", conferencia impartida en el Instituto Tecnológico Autónomo de México, México, D.F., 25 de abril de 2003.
- SOLARTE PAZOS, Leonardo, 2002, "La evaluación de programas sociales en el Estado neoliberal", *VII Congreso Internacional del CLAD sobre la Reforma del Estado y la Administración Pública*, Lisboa, 8-11 de octubre.
- SPINK, Peter, 2001, "Modernización de la Administración Pública en América Latina, evaluación de los recientes cambios en la gerencia de los ministerios y de las agencias reguladoras", en CLAD, Serie Estudios de Casos, *Nueva Gestión Pública y regulación en América latina. Balances y desafíos*, [www.clad.org.ve](http://www.clad.org.ve), consulta realizada el 4 de octubre de 2004.
- SPIZZ, Pierre, 2002, "Learning and/or measuring performance? Competing priorities within the UN system", ponencia presentada en la *V Conferencia Europea de Evaluación*, Sevilla, [www.europeanevaluation.org](http://www.europeanevaluation.org)
- STUFFLEBEAN, Daniel, 1974, "Meta-evaluation", en *Ocasional Paper* núm. 3, [www.wmich.edu](http://www.wmich.edu), consulta efectuada el 2 de mayo de 2003.
- SZÉKELY PARDO, Miguel, 1990, *Identificación de la pobreza global y rural en México, un primer paso para incrementar el grado de acierto de los programas públicos*, tesis para obtener el grado de maestro en políticas públicas, ITAM.
- , diciembre de 1993, "El Programa Nacional de Solidaridad en México", *documento de trabajo* núm. 384, OIT.
- , 2003, "Lo que dicen los pobres", *Cuadernos de Desarrollo Humano*, núm. 13, México, Sedesol.
- THOENIG, Jean-Claude, junio 1996, Conferencia sobre Evaluación, CIDE A.C., México.
- , noviembre 2002, "Upgrading knowledge in public policy and administration", ponencia presentada en el Seminario Internacional *El Estado el Arte en Administración y Políticas Públicas*, México, Colegio de México.
- , 2003, presentación titulada "Evaluation of National Public Policies. Learning from experience", *documento en Power Point*.
- TORRES, Paul, 2000, "Se former à l'évaluation de l'aide au développement du territoire", en Comunicación del Grupo de Apoyo al Desarrollo para el *Coloquio 2000 de la Sociedad Francesa de Evaluación*, [www.evalua.com](http://www.evalua.com), consulta realizada el 2 de mayo de 2003.
- TOULEMONDE, Jacques, 2000, "Key lessons from the European-wide evaluation practice in the field of social and economics development",

- ponencia presentada en la *IV Conferencia Europea de Evaluación*, Lausanne, [www.europeanevaluation.org](http://www.europeanevaluation.org).
- , 2002, “Assuring the quality of evaluations, lessons from European experience”, ponencia presentada en la *V Conferencia Europea de Evaluación*, Sevilla, [www.europeanevaluation.org](http://www.europeanevaluation.org).
- U.A. DE CHAPINGO, 2003, “Evaluación externa de los programas del Fondo de Fomento para las Artesanías, Fonart)”, en *CD entregado por el IFAI*.
- UAM-A, 2003, “Evaluación externa del Programa Identidad Jurídica. Informe final”, en *CD entregado por el IFAI*.
- UAM-X, 2003, “Informe definitivo 2002. Programa Promoción de las Culturas Indígenas del Instituto Nacional Indigenista”, en *CD entregado por el IFAI*.
- UAM-X y Norlatina Consultores, 2003, “Evaluación externa del programa de Empleo Temporal 2001-2002”, en *CD entregado por el IFAI*.
- U.A. DE TLAXCALA, 2003, “Programas Estatales por Demanda”, en *CD entregado por el IFAI*.
- , 2003, “Programas Jóvenes por México”, en *CD entregado por el IFAI*.
- , 2003, “Programa Mujeres Jefas de Familia”, en *CD entregado por el IFAI*.
- , 2003, “Informe final de la evaluación externa del Programa Iniciativa Ciudadana 3X1”, en *CD entregado por el IFAI*.
- UVALLE, Ricardo, octubre de 2002, ponencia “Los fundamentos institucionales de la gestión pública”, en *CD del VII Congreso Internacional del CLAD sobre la Reforma del Estado y la Administración Pública*, Lisboa, Portugal.
- , octubre de 2003, ponencia “La democratización de la Administración Pública”, en *CD del VIII Congreso Internacional del CLAD sobre la Reforma del Estado y la Administración Pública*, Panamá, Panamá.
- VARIOS AUTORES, 2000, “Indicadores de gestión para las entidades públicas”, *Documento núm. 16*, Madrid, AECA.
- WAISSBLUTH, Mario, 2004, *Seminario de Gerencia Pública*, impartido en la sede del BID, EUA.

# Índice

PRESENTACIÓN	
Luis F. Aguilar . . . . .	7
INTRODUCCIÓN GENERAL . . . . .	13

## *Primera parte*

### EVALUACIÓN DE POLÍTICAS Y PROGRAMAS PÚBLICOS

#### *Capítulo 1*

PAPEL DE LA EVALUACIÓN EN EL CONTEXTO DE LA ADMINISTRACIÓN PÚBLICA . . . . .	23
Introducción . . . . .	23
La administración pública tradicional . . . . .	23
El estudio de las políticas públicas . . . . .	24
El New Public Management . . . . .	30
Aportes sociológicos . . . . .	36
Enfoques emergentes . . . . .	39
Comentario final . . . . .	41

#### *Capítulo 2*

LA EVALUACIÓN DESDE EL PUNTO DE VISTA TEÓRICO . . . . .	43
Introducción . . . . .	43
Concepto de evaluación . . . . .	43
Justificación de la evaluación . . . . .	45
Tipos de evaluación . . . . .	47
Actitudes y enfoques . . . . .	51
Contexto político-organizativo de la evaluación . . . . .	58
Comentario final . . . . .	60

*Capítulo 3*

MÉTODOS Y TÉCNICAS DE EVALUACIÓN . . . . .	61
Introducción . . . . .	61
Debate sobre el método . . . . .	61
Experimentación y métodos cuantitativos . . . . .	63
Métodos cualitativos . . . . .	70
Comentario final . . . . .	76

*Capítulo 4*

EXPERIENCIAS DE EVALUACIÓN EN ORGANISMOS INTERNACIONALES Y PAÍSES DESARROLLADOS . . . . .	79
Introducción . . . . .	79
Ámbito internacional: ¿control o evaluación participativa? . . . . .	79
Países desarrollados . . . . .	83
EUA: los dilemas de la experimentación . . . . .	85
Reino Unido: un efectivo trabajo de convencimiento y capacitación . . . . .	88
Canadá: ¿evaluación real o mero ritual? . . . . .	91
Francia: en búsqueda de un modelo adaptado a su idiosincrasia . . . . .	94
Comentarios finales . . . . .	103

*Capítulo 5*

DESARROLLOS EMPÍRICOS DE LA EVALUACIÓN EN AMÉRICA LATINA . . . . .	105
Introducción . . . . .	105
América Latina: ejemplos escasos y parciales . . . . .	105
México: una sistematización largamente postergada . . . . .	115
Comentarios finales . . . . .	125

*Capítulo 6*

META-EVALUACIÓN DE LAS EVALUACIONES . . . . .	129
Introducción . . . . .	129
Dificultades para implantar procesos de evaluación . . . . .	129
Conceptos e instrumentos de metaevaluación . . . . .	132
Criterios de metaevaluación . . . . .	139
Tipos de metaevaluación . . . . .	140
Experiencias de metaevaluación . . . . .	141
Propuesta de metaevaluación a aplicar en México . . . . .	146
Comentarios finales . . . . .	148

*Segunda parte*  
**LA EVALUACIÓN DE POLÍTICAS Y  
PROGRAMAS MEXICANOS DE DESARROLLO SOCIAL**

*Capítulo 7*

POLÍTICAS Y PROGRAMAS DE DESARROLLO SOCIAL . . . . .	153
Introducción . . . . .	153
Política social y pobreza . . . . .	153
Programas aplicados en sexenios previos . . . . .	156
Plan Nacional de Desarrollo 2001-2006 . . . . .	159
Programa Nacional de Desarrollo Social 2001-2006 . . . . .	164
Estrategias Contigo y Microrregiones . . . . .	165
Desarrollo de capacidades . . . . .	168
Desarrollo del patrimonio . . . . .	171
Generación de oportunidades . . . . .	172
Otros programas . . . . .	174
Comentarios finales . . . . .	175

*Capítulo 8*

PRIMEROS ESFUERZOS DE EVALUACIÓN DE PROGRAMAS . . . . .	177
Introducción . . . . .	177
Particularidades de la evaluación de políticas sociales . . . . .	177
Evaluaciones de Pronasol . . . . .	179
Evaluaciones de Progresa 1998-2000 . . . . .	184
Auditorías de desempeño de Progresa 1999-2000 . . . . .	188
Evaluación y auditoría de desempeño de Progresa 2001 . . . . .	189
Comentarios finales . . . . .	193

*Capítulo 9*

MARCO JURÍDICO DE LA EVALUACIÓN DE PROGRAMAS . . . . .	195
Introducción . . . . .	195
Control y evaluación al interior de la administración pública federal . . . . .	195
Planeación y organización de la evaluación en Sedesol . . . . .	199
Papel del Congreso . . . . .	202
Regulación de la contratación de evaluaciones externas . . . . .	208
Disposiciones generales . . . . .	208
Bases y términos de referencia . . . . .	215



Ley Federal de Transparencia y Acceso a la Información Pública Gubernamental . . . . .	216
Ley General de Desarrollo Social . . . . .	218
Comentarios finales . . . . .	219

### *Capítulo 10*

ORGANIZACIÓN Y GESTIÓN DEL SISTEMA DE EVALUACIONES EXTERNAS . . . . .	221
Introducción . . . . .	221
Contexto político y cultural . . . . .	221
Tipos de evaluación planeados . . . . .	223
Decisión de realizar evaluaciones externas . . . . .	226
Organización y gestión del sistema . . . . .	227
Organización y gestión formal . . . . .	228
Funcionamiento real del sistema . . . . .	229
Recursos para la realización de evaluaciones externas . . . . .	239
Recursos financieros . . . . .	239
Recursos humanos . . . . .	243
Otros recursos . . . . .	249
Comentarios finales . . . . .	249

### *Capítulo 11*

MÉTODOS, CALIDAD Y COSTOS DE LAS EVALUACIONES EXTERNAS . . . . .	251
Introducción . . . . .	251
Métodos y procesos de las evaluaciones externas . . . . .	252
Fuentes de información y análisis de datos . . . . .	252
Pertinencia y diseño del programa . . . . .	258
Recursos utilizados . . . . .	262
Análisis de procesos . . . . .	265
Medición de cobertura . . . . .	269
Resultados alcanzados . . . . .	272
Impactos provocados . . . . .	274
Participación y satisfacción social . . . . .	277
Escenarios prospectivos . . . . .	280
Valoración global del programa . . . . .	281
Utilización de resultados y recomendaciones de evaluaciones previas . . . . .	281
Informes finales . . . . .	283

Costo/calidad de los resultados de las evaluaciones externas . . . . .	283
Juicio valorativo de la metaevaluación transversal . . . . .	286
Evolución de algunas evaluaciones seleccionadas . . . . .	288
Comentarios finales . . . . .	291

### *Capítulo 12*

#### IMPACTOS Y PERSPECTIVAS DE LA EVALUACIÓN

DE PROGRAMAS DE DESARROLLO SOCIAL . . . . .	295
Introducción . . . . .	295
Caracterización de las evaluaciones externas analizadas . . . . .	295
Impacto de las evaluaciones . . . . .	298
Análisis FODA del sistema de evaluación . . . . .	299
Árbol de variables y problemas identificados . . . . .	302
Propuestas . . . . .	312
Escenarios prospectivos . . . . .	318
Conclusiones . . . . .	321

#### CONSIDERACIONES FINALES . . . . . 325

Las tendencias en el debate teórico sobre evaluación de programas . . . . .	325
El reciente informe del Banco Mundial sobre la pobreza y las estrategias del gobierno mexicano . . . . .	326
Expectativas de contribución al debate actual . . . . .	328

### *Anexo 1*

Normas para la evaluación de programas del Joint Committee on Standards for Educational Evaluation Segunda edición, 1994 . . . . .	337
Normas de utilidad . . . . .	331
Normas de factibilidad/viabilidad . . . . .	332
Normas de corrección/legitimidad . . . . .	332
Normas de precisión . . . . .	333

### *Anexo 2*

Principales funciones de las secretarías de Estado vinculadas a los procesos de evaluación . . . . .	335
Secretaría de Hacienda y Crédito Público (SHCP) . . . . .	335
Secretaría de la Función Pública (SFP) . . . . .	335

SIGLAS Y ACRÓNIMOS . . . . .	339
BIBLIOGRAFÍA . . . . .	343
Libros y revistas . . . . .	343
Documentos, ponencias, periódicos, tesis y consulta a Internet . . . . .	356

*La evaluación de políticas y programas públicos. El caso de los programas de desarrollo social en México, se terminó de imprimir en la ciudad de México durante el mes de agosto del año 2006. La edición, en papel de 75 gramos, consta de 1,500 ejemplares más sobrantes para reposición y estuvo al cuidado de la oficina litotipográfica de la casa editora.*



El crecimiento indiscriminado de la burocracia en México llevó a que en las últimas dos décadas la administración pública replanteara sus esquemas e hiciera nuevas propuestas que incluyeran tres características fundamentales: eficacia, economía y eficiencia. Las políticas públicas como fenómenos sociales, administrativos y políticos específicos son resultado de un proceso de sucesivas tomas de posición en respuesta a los problemas públicos que aquejan a la sociedad civil; de esta manera los gobiernos son los principales actores que disponen de recursos suficientes para ofrecer bienes y servicios para satisfacer las necesidades sociales. Este trabajo pondera la evaluación como herramienta de buena gestión, tanto para aplicación de incentivos positivos y negativos, como para la transparencia contable y la rendición de cuentas, que son insumos básicos de la gobernancia con participación; así la traducción de los objetivos estratégicos del gobierno en metas operativas y la verificación de su cumplimiento constituyen la verdadera esencia de la gerencia pública.

La evaluación de políticas



9 789707 018242

Miguel Ángel  
Porrua



**CONOCER  
PARA DECIDIR**  
EN APOYO A LA  
INVESTIGACIÓN  
ACADÉMICA