



México: Evaluación de políticas rurales. Tendencias teóricas y lecciones de la experiencia

**Grupo Interagencial de Desarrollo
Rural – México**

México, marzo de 2009



Distr.
LIMITADA

LC/MEX/L.901
25 de marzo de 2009

ORIGINAL: ESPAÑOL

**MÉXICO: EVALUACIÓN DE POLÍTICAS RURALES.
TENDENCIAS TEÓRICAS Y LECCIONES
DE LA EXPERIENCIA**

ÍNDICE

	<u>Página</u>
RESUMEN EJECUTIVO	1
PRESENTACIÓN.....	3
I. EVALUACIÓN DE POLÍTICAS RURALES: TENDENCIAS TEÓRICAS Y ENFOQUES INSTITUCIONALES.....	5
1. Importancia de la evaluación.....	5
2. Evolución y tendencias teóricas actuales	7
3. Evolución reciente de los trabajos sobre evaluación de políticas	13
4. Enfoques institucionales de la evaluación de políticas y proyectos...	15
5. Resumen	19
II. APRENDIZAJES DE LA EVALUACIÓN DE PROGRAMAS DE FOMENTO: EL CASO DE ALIANZA PARA EL CAMPO.....	21
1. Introducción	21
2. La Alianza para el Campo.....	21
3. La evaluación de Alianza para el Campo.....	22
III. EVALUACIÓN DE POLÍTICAS PÚBLICAS PARA EL DESARROLLO RURAL SUSTENTABLE: EL CASO DE MÉXICO	33
1. Sistemas públicos de información.....	34
2. Instrumentos de evaluación: Encuestas.....	34
3. Estimaciones.....	35
BIBLIOGRAFÍA	39
<u>Anexo 1:</u> Enlaces sobre evaluación	43

RESUMEN EJECUTIVO

En este trabajo se destaca la importancia de evaluar programas, proyectos y políticas, en particular rurales, a fin de mejorar su eficiencia, efectividad, el rendimiento de cuentas, la asignación de recursos, y en especial para la creación de una política pública. Asimismo, se resalta la contribución de la evaluación para un mejor diseño en la formulación de estrategias y políticas rurales, que desemboquen en un mayor impacto.

En primer lugar, en este estudio se revisan los fundamentos teóricos del proceso de evaluación y sus ramas principales de desarrollo, así como los métodos, el uso de la evaluación y la valoración que el evaluador o los interesados hacen del proceso. En el caso de los métodos, se comenta sobre la evaluación orientada por objetivos, la dirigida por teoría y las diseñadas para cada programa, el diseño experimental y cuasi experimental, la investigación evaluativa, la consideración especial por el contexto y los *stakeholders*.

Con respecto a la valoración, se examina la capacidad del evaluador para juzgar el programa, política o proyecto en su calidad de experto. También se estudia la evaluación por adversarios que complementen adecuadamente el juicio, así como la consideración de múltiples contextos y perspectivas de los interesados en ésta. Asimismo, se analiza el hecho de que la evaluación no es neutral y tiene que atender las necesidades y los intereses de los menos favorecidos.

En cuanto al uso de los resultados de la evaluación, se comenta que ha sido pensado para que los *stakeholders* y los administradores de programas tomen decisiones. Se sostiene que los evaluadores tienen que promover la evaluación entre estos usuarios e identificar nuevos.

En el trabajo se resume el aporte valioso que han hecho instituciones como el Banco Mundial, la FAO, la CEPAL y el IICA en cuanto a metodologías, criterios, realización de evaluaciones, y lecciones aprendidas durante los últimos 40 años.

En este análisis se examinan las evaluaciones realizadas al principal programa de fomento productivo rural de México: Alianza para el Campo, los últimos 10 años por medio de un proyecto conjunto del SAGARPA y la FAO. Se comentan los retos encontrados para las evaluaciones, las metodologías empleadas, los logros obtenidos, las limitaciones y las lecciones aprendidas en la realización de los procesos de evaluación.

De esta forma, se presentan lecciones derivadas de la evaluación sumativa, la importancia de contar con información contrafactual, la identificación de indicadores básicos que pueden ser utilizados pese a los cambios de programa, y la conveniencia de prever la evaluación desde el inicio para definir indicadores de desempeño y darles seguimiento. Finalmente, se busca considerar los tiempos de maduración de las inversiones al definir la periodicidad de las evaluaciones.

También se concluye que desde el diseño del programa se deben identificar las necesidades de información para los indicadores. Dadas las limitaciones de información, los estudios de caso o grupos focales resultan valiosos para las evaluaciones en el medio rural. A fin de evitar la dispersión, se recomienda acotar la recolección de datos a los verdaderamente necesarios para la evaluación. La participación de los operadores en el diseño del programa es valiosa, ya que éstos serán sus principales usuarios.

Generalmente, se subestima la evaluación formativa o de procesos. Sin embargo, de las evaluaciones de Alianza para el Campo se derivaron varias *lecciones* de utilidad para futuras valoraciones de programas rurales. Si se definen previamente los indicadores, que retroalimentan a los que toman las decisiones, la evaluación debe estar incorporada a la gestión de los programas. Una deficiente instrumentación puede ser corregida por medidas derivadas de la evaluación de procesos, ya que ésta tiene mayor utilidad si se extiende a un análisis de políticas públicas.

En el caso de programas descentralizados, la evaluación de procesos debe incluir los diferentes niveles mediante enfoques, instrumentos y metodologías propias. Con este tipo de evaluación se deben identificar los procesos con mayor incidencia en los impactos del programa y se concluye que la observación directa de los evaluadores y las entrevistas abiertas con operadores y beneficiarios son fuente importante de información cualitativa, que se debe complementar con el análisis documental y de encuestas.

En México el IICA ha trabajado en la evaluación del fortalecimiento institucional para el desarrollo rural, al desarrollar el Sistema de Monitoreo y Evaluación del Desarrollo Rural Municipal, SIMOM. De la experiencia del diseño e implementación de este sistema de monitoreo, se han recogido experiencias que contribuyen a la valoración de la evaluación como un instrumento de gerencia pública y de control social. Así, las *lecciones* aprendidas son las siguientes:

La cultura de evaluación debe anteceder al desarrollo técnico de instrumentos de medición. La evaluación, como un mecanismo interno de gestión, tiene grandes potencialidades para alertar anticipadamente en la reconducción de los procesos de gestión de la política. La evaluación de procesos, como la aplicada por el SIMOM, se complementa con la evaluación de resultados e impactos.

Asimismo, el acceso y uso de la información de la evaluación es un desafío. Se exige en destacar el tema de la comunicación y la difusión de los resultados de la evaluación, lo que incluye un verdadero proceso pedagógico. Los sistemas de evaluación requieren de agentes técnicos de soporte, como el caso del SIAP para el SIMOM. Y como la evaluación de procesos no puede ser una evaluación externa, *ex post*, debe ser entendida como un proceso de medición y observación de la operación, en tiempo real, por los actores mismos y con la aplicación inmediata de correctivos.

PRESENTACIÓN

La profunda crisis global actual, de astringencia de recursos y gran volatilidad de mercados, entre ellos los agropecuarios, exige un uso muy eficiente de los escasos recursos públicos y de cooperación internacional. Asimismo, la restringida disponibilidad de tierras, aguas, energía y fertilizantes demandan un uso eficiente de estos recursos. En este marco, la evaluación de políticas y proyectos públicos rurales puede contribuir significativamente a eficientar el uso de los recursos.

En los últimos años, la evaluación y seguimiento de políticas rurales han respondido a varios temas prioritarios. En especial, se ha buscado hacer más eficiente el desempeño de las instituciones públicas y mostrar a la cooperación bilateral y multilateral los impactos de su accionar, así como a responder al “desencanto” con los logros en el desarrollo rural. También se ha tenido como propósito lograr que los nuevos temas prioritarios del desarrollo (género, cambio climático y sanidad e inocuidad de alimentos) se atiendan con efectos positivos.

A pesar del amplio desarrollo teórico y de las aplicaciones logradas en materia de evaluación y seguimiento de políticas y programas, existe un gran campo de aplicación en los países de la Subregión Norte y considerables posibilidades de cooperación entre los mismos.

Por lo anterior, con el Grupo Interagencial de Desarrollo Rural–México se desea contribuir al diálogo y ampliación de los trabajos de evaluación. Para ello se realizó un conjunto de reflexiones acerca del tema de evaluación plasmadas en el presente estudio acerca de las tendencias teóricas actuales, el aprendizaje de la experiencia de evaluación de Alianza para el Campo y las consideraciones sobre el sistema de monitoreo y Evaluación del Desarrollo Rural Municipal en México.

El análisis cubre tres temas. En el capítulo I se comenta acerca de la importancia de la evaluación para los gobiernos, las sociedades y la cooperación internacional. Se revisan los fundamentos teóricos y algunos desarrollos recientes de la teoría de la evaluación y se describe brevemente la evolución de los trabajos empíricos sobre evaluación de políticas y los principales enfoques y aportes institucionales de la FAO, el BIRF, la CEPAL y el IICA.

El propósito del capítulo II es compartir los aprendizajes derivados de la experiencia de evaluación de Alianza para el Campo. Así, se examinan las evaluaciones realizadas al principal programa de fomento productivo rural de México durante los últimos 10 años y se comentan, de manera sintética, las metodologías empleadas, sus logros, limitaciones y algunas importantes lecciones aprendidas en los procesos de evaluación.

En el capítulo III se analizan los trabajos del IICA en el Sistema de Monitoreo y Evaluación del Desarrollo Rural Municipal de México y se enfatizan los aspectos instrumentales —la definición y elaboración de indicadores— y metodológicos del seguimiento a nivel local.

Este estudio *México: Evaluación de políticas rurales. Tendencias teóricas y lecciones de la experiencia* se enmarca en las tareas del Grupo Interagencial de Desarrollo Rural (GIADR)¹ de diálogo con la SAGARPA y otros actores públicos y privados interesados en el desarrollo rural de México.

¹ El Grupo fue creado en la Asamblea Anual de los Gobernadores del BID en Nueva Orleans (marzo de 2000), con el propósito de complementar la acción de las agencias que lo integran mediante la generación y difusión de información relevante para el desarrollo de las economías rurales de los países de América Latina y el Caribe; la identificación de instrumentos, de nuevos enfoques y de experiencias exitosas de desarrollo rural, y la implementación de proyectos conjuntos de capacitación, asistencia técnica e inversión a nivel nacional y regional. A nivel latinoamericano, dicho grupo está formado por ocho agencias de cooperación técnica y financiamiento para el desarrollo: Agencia Alemana de Cooperación Técnica (GTZ), Agencia de los Estados Unidos para el Desarrollo Internacional (USAID), Banco Interamericano de Desarrollo (BID), Banco Mundial (BM), Comisión Económica para América Latina y el Caribe (CEPAL), Fondo Internacional de Desarrollo Agrícola (FIDA), Instituto Interamericano de Cooperación a la Agricultura (IICA) y la Organización de las Naciones Unidas para la Agricultura y la Alimentación (FAO).

I. EVALUACIÓN DE POLÍTICAS RURALES: TENDENCIAS TEÓRICAS Y ENFOQUES INSTITUCIONALES

En este capítulo se comenta la importancia de la evaluación para los gobiernos, las sociedades y la cooperación internacional. Se revisan los fundamentos teóricos y algunos desarrollos recientes de la teoría de la evaluación y se describe brevemente la evolución de los trabajos empíricos sobre evaluación de políticas y los principales enfoques y aportes institucionales, de la FAO, el BIRF, la CEPAL y el IICA.

1. Importancia de la evaluación

En general la evaluación es una disciplina fundamental que se relaciona con muchas actividades humanas con la intención de mejorar las condiciones y resultados del ámbito evaluado. Se aplica en múltiples espacios de la producción de bienes y servicios y sus importantes implicaciones inciden en el mantenimiento, mejoramiento y en la protección de los ciudadanos. En la evaluación se examina la efectividad con relación al costo, la confiabilidad, la eficiencia, la seguridad, la facilidad de uso o la confiabilidad del objeto evaluado, sea un bien, servicio, proyecto, programa o política (Stufflebeam y Shinkfield, 2007).

La evaluación es importante para las sociedades, ya que les confirma el valor, la importancia, la mejora del ente evaluado —y de cómo y cuándo ésta sucedió—, la acreditación, el rendimiento de cuentas, dar cuenta (*accountability*) y, cuando es necesario, proporciona bases para terminar con políticas o programas malos. La evaluación por sí sola no es suficiente para garantizar servicios o bienes de alta calidad; se requiere de responsables en las tomas de decisiones que implementen los resultados de la evaluación. Así, las sociedades ameritan no sólo de evaluadores competentes, sino de responsables que decidan qué hacer con los resultados de las evaluaciones (Stufflebeam y Shinkfield, 2007).

En los últimos años la evaluación y el seguimiento de políticas, programas y proyectos rurales han crecido en importancia debido a la necesidad de hacer un uso eficiente y eficaz de los recursos públicos y de la cooperación internacional, así como de su aplicación a nuevos temas prioritarios.

a) En primer lugar, se constata la necesidad de mejorar el desempeño público en los países, y la evaluación de políticas y programas rurales puede contribuir a este propósito. Los logros limitados de la acción pública en el medio rural no siempre positiva correlación entre gasto público y aumentos en ingresos rurales, productividad y competitividad externa, así como las expectativas de los ciudadanos por mayores y mejores resultados del gasto, dan lugar a una mayor demanda de evaluaciones de la acción pública en el medio rural.

Se busca lograr mayor efectividad de los programas gubernamentales; conocer sus logros seguros y sus resultados, así como determinar si realmente llegan a las poblaciones objetivo y si tienen un impacto en las variables de política rural.

b) En la misma forma los ciudadanos de los países donantes y de los organismos financieros y cooperantes exigen mayor control en el uso de sus recursos para la cooperación internacional. En varios casos, durante la década pasada, se demostró cierto cansancio con respecto a la cooperación internacional para la agricultura, lo que llevó a su disminución. Por ello, cada vez es más frecuente el uso de evaluaciones, en particular, las de impacto y monitoreo en los proyectos de cooperación internacional.

c) La insatisfacción por el pobre desarrollo rural es una tercera fuente que impulsa el uso creciente de evaluaciones de políticas y proyectos. Hay un cierto agotamiento de las estrategias y políticas rurales, por ejemplo: fuertes apoyos a precios y comercialización; intervención estatal en la producción y distribución por medio de entes públicos; desregulación y supresión de organismos estatales de apoyo y apertura comercial. Los logros en cuanto a crecimiento agropecuario, reducción significativa de la pobreza y de la desigualdad, y disminución de las migraciones rurales han sido limitados.

Lo anterior requiere de más y mejores evaluaciones —y seguimiento de políticas y programas— para identificar nuevos y más eficaces caminos al desarrollo rural, conocer mejor las experiencias exitosas y sus posibilidades de réplica a mayor escala, e identificar las restricciones institucionales y técnicas principales que coartan la eficiencia de las políticas públicas.

d) En cuarto lugar, en la actualidad han surgido necesidades adicionales de evaluación y seguimiento derivadas de los nuevos temas prioritarios del desarrollo rural, como género y mejor inserción de la mujer; la atención a los efectos del cambio climático sobre el medio rural y el nuevo énfasis en sanidad e inocuidad de alimentos. La evaluación de las nuevas políticas permite ver su eficacia al tratar estos nuevos temas y definir mejores instrumentos.

e) La lucha por reducir la pobreza y el cumplimiento de las metas del milenio sobre este importante desafío requieren de exámenes periódicos a fin de revisar el avance en las metas y evaluaciones de impacto para conocer los logros y las estrategias y políticas más eficaces.

f) En varios países del medio rural, los recursos públicos dedicados a la agricultura son muy escasos. Se tiende a limitarlos fiscalmente, a generarles déficit públicos, y a contraer el gasto público agropecuario en proporción al total, de por sí reducido y que no es proporcional al aporte que hace el sector a la economía. Por otra parte, se constata que el gasto en subsidios privados es alto y por lo tanto poco eficiente y excluye de apoyos a una parte considerable de productores.

Estas restricciones e ineficiencias en el gasto público hacen que la evaluación y seguimiento del gasto sea muy importante, lo que nuevamente evidencia la importancia y el potencial de los sistemas de evaluación y monitoreo de políticas, programas y proyectos.

g) Para concluir, las recientes tendencias en las cotizaciones internacionales de alimentos constituyen un nuevo ámbito de acción que requiere de la evaluación de políticas, que representan tanto un desafío a la seguridad alimentaria como una oportunidad para elevar la oferta de manera sostenida, competir internacionalmente y mejorar los ingresos de los

productores, en especial de los pequeños. Para ello, las políticas agropecuaria y rural deben mejorar su impacto y las evaluaciones contribuirán efectivamente a ello.

En síntesis, la importancia de los procesos de evaluación y seguimiento radica en que se provea de información valiosa sobre el desempeño de las políticas, programas y proyectos gubernamentales, en la que se identifiquen qué es lo que funciona, qué no y por qué. Con la evaluación y el seguimiento también se suministra información del trabajo del gobierno o de ministerios o agencias públicas, en particular de sus directivos y funcionarios, y además se proporciona información acerca del desempeño de los donantes que apoyan el trabajo de los gobiernos (Banco Mundial, 2007). Como valiosa herramienta para la toma de decisiones, eleva la eficacia y eficiencia de la política pública rural.

2. Evolución y tendencias teóricas actuales

La teoría, métodos y uso de la evaluación han tenido un gran desarrollo en los últimos 50 años. Una muy buena síntesis de la evolución del pensamiento sobre evaluación se encuentra en el trabajo de Alkin y Christie, *An Evaluation Theory Tree* (Alkin, 2004). A continuación se presenta una selección de las principales ideas contenidas en ese trabajo.² El análisis teórico de la evaluación mostrado se concentra en la reflexión sobre los *métodos* a utilizar, la *valoración* implícita en las evaluaciones y el *uso* que se da a los resultados.

Los aportes teóricos en el ámbito de los métodos han incluido temas como la *evaluación orientada por objetivos o evaluación referida a objetivos*; en el diseño experimental y cuasi experimental para la investigación, la aleatoriedad es lo más importante; las nociones de *validez interna* y de *validez externa*; el criterio administrativo para determinar la utilidad del estudio; la importancia de considerar el contexto y de usar diferentes diseños y métodos para una evaluación apropiada, la participación de los *stakeholders*, y el concepto de evaluación dirigida por teoría y evaluación comprensiva.

En materia de valoración, diversos teóricos enfatizan la importancia de hacer juicios en los procesos de evaluación, aunque a través de diferentes maneras y roles para los evaluadores. El evaluador debe ser un experto en la materia que evalúa y selecciona lo que describe y critica. Se propone la *evaluación por adversarios*; algunos teóricos consideran que no hay verdadero valor posible, el conocimiento está sujeto al contexto y la perspectiva de los *stakeholders* es parte integral de la evaluación. La evaluación nunca es neutral, la justicia, la equidad y lo correcto la modifican.

En cuanto al uso, algunos teóricos sostienen que las evaluaciones se deben diseñar como herramientas para asistir a los *stakeholders* en la toma de decisiones. Otros consideran que las evaluaciones tienen que apoyar a los administradores de programas y *policy makers*, para asignar recursos a los programas que sirven mejor a los clientes.

Hay dos fundamentos básicos de las teorías de evaluación: la rendición de cuentas (*accountability*) y la investigación social. La intención de rendición de cuentas implica una

² Ésta es una traducción libre realizada por el autor de este capítulo.

necesidad de evaluación. La preocupación por emplear un conjunto sistemático y justificable de métodos para determinar la rendición de cuentas lleva a la necesidad de fundamentarse en la investigación social. Muchos teóricos han reconocido la importancia de la rendición de cuentas, al identificar sus dimensiones. En la primera fase se reporta y se proporciona una descripción. En la segunda se explica o se justifica el análisis. En la tercera se obliga a los responsables a responder a las observaciones de la evaluación. Además, se habla del proceso de rendición de cuentas en relación con las metas, el proceso y los resultados o efectos: ¿han sido razonables y apropiadas las metas y los procesos?; ¿en qué grado se han cumplido las metas? Esto último es el tema de la mayor parte de los esfuerzos de evaluación.

Por su parte, la investigación social,³ fundamento de la teoría de la evaluación, busca dar respuesta a la pregunta central, ¿por qué personas en grupos sociales actúan de cierta forma? Esta búsqueda avanza paralelamente a la permanente cuestión de cuáles son los métodos apropiados para la investigación social y las evaluaciones. La discusión acerca de la aplicabilidad y la pertenencia de los métodos utilizados para los estudios sociales continúa en la actualidad. Así también, se mantienen preguntas propias a la teoría de la evaluación como ¿cuál es la relación entre teoría y observación?; ¿deberían los científicos sociales tener una opinión moral hacia los grupos o individuos que estudian?; ¿es ésta una opinión apropiada? o ¿podría comprometer su objetividad?

Alkin y Christie en su estudio clasifican los desarrollos teóricos en evaluación de acuerdo con tres grandes ramas del pensamiento: a) Métodos. Es la evaluación como investigación, se enfatiza en los métodos de evaluación; b) Valoración. Establece el papel vital que debe desempeñar el evaluador en valorar al ente evaluado —proyecto, política o programa— y dar valor a los datos, y c) Uso. Se preocupa por la manera en que la información —resultados— de la evaluación será utilizada y en aquellos que la usarán.

A continuación se presenta, a partir de una traducción libre, un resumen de las ideas del valioso estudio de Alkin y Christie (2004). Se hace un *recorrido* del *árbol de* Teoría de la Evaluación a través de los principales desarrollos, de manera temática, y no por autores, de las tres grandes ramas teóricas de la evaluación, métodos, valoración y usos.

a) Métodos

La mayoría de los teóricos de la evaluación plantean preocupaciones metodológicas y ven a la *investigación* como la génesis de la evaluación de programas. Ralph Tyler, uno de los fundadores de la teoría de la evaluación se concentró en la especificación de objetivos y la medición de resultados, *evaluación orientada por objetivos o evaluación referida a objetivos*. Tyler (1942) acuñó el concepto de evaluación educacional para describir su proceso de comparación de “objetivo —resultados bien definidos que se persiguen— con bien medidos resultados realmente alcanzados”.

³ La investigación social se entiende como el estudio sistemático del comportamiento de grupos de individuos en varias clases de escenarios o ambientes sociales, por una variedad de métodos.

Por muchas décadas la evaluación orientada por objetivos ha sido utilizada en la educación.⁴ Autores como Donald Campell con J. Stanley, en su trabajo sobre el diseño experimental y cuasi experimental para la investigación, explicaron las condiciones necesarias para un verdadero estudio experimental, y precisaron a la aleatoriedad como la más importante. Asimismo, definieron dos nociones clave: la *validez interna*, es decir, el grado en el que un experimento es controlado de manera apropiada; y la *validez externa* o el grado de amplitud en que los resultados de un experimento son aplicables.

Otro aporte de estos autores fue reconocer (Suchman, 1967) que los experimentos no son perfectos —podrían ser no alcanzables o no deseables— y por ello no deberían usarse en muchos casos. Como una alternativa plantearon los cuasi experimentos, que incluyen algún tipo de intervención o tratamiento y proveen una comparación, pero carecen del grado de control que tienen los experimentos verdaderos. Así como la aleatoriedad es la clave de los experimentos verdaderos, la falta de aleatoriedad es una característica que define a los cuasi experimentos.

Edward Suchman (1967) vio la relevancia de estos métodos de investigación para la evaluación y distinguió entre el concepto de evaluación en sentido común, como hacer juicios de valor, y la investigación evaluativa que utiliza métodos y técnicas de investigación científica. Varias guías gubernamentales de evaluación tienen como gran deficiencia la ausencia de un diseño experimental.

Suchman sostiene que el evaluador debe tener, además del criterio científico reconocido, criterio administrativo para determinar la utilidad del estudio y si mediante un programa o proceso se logra algún resultado deseado. El éxito de la evaluación dependerá, en gran medida, de la utilidad que tenga para que el administrador mejore los servicios. Suchman estableció cinco categorías de evaluación: a) *esfuerzo*, cantidad y calidad de actividad; b) *desempeño*, mide los resultados del esfuerzo; c) *adecuación del desempeño*, grado en que el desempeño es adecuado a la cantidad de necesidad; d) *eficiencia*, examen de sendas o métodos alternativos en costos monetarios y de trabajo, y e) *procesos*, cómo y porqué un programa/política funciona o no.

Thomas Cook (2005) enfatizó los métodos cuasi experimentales y se encargó también de los factores contextuales y cómo éstos pueden afectar a la evaluación. Para realizar una evaluación apropiada, Cook acentuó la importancia de usar diferentes diseños y métodos que tomen en consideración el *contexto* de cada evaluación, en lugar de usar el mismo conjunto de métodos y diseños para todas las evaluaciones. Esto, además de ser un ataque directo al diseño experimental en las evaluaciones, reconoce también la importancia de involucrar a los interesados, en especial a los responsables en la toma de decisión —*stakeholders*—, al definir las preguntas que contendrá la evaluación, lo que evita que los resultados sean operacionalmente vagos o de limitada importancia para muchos usuarios potenciales.

Para algunos teóricos como Robert Boruca (1994), la evaluación es como una investigación científica convencional, que se define por experimentos de campo aleatorios controlados. Asimismo, Peter Rossi contribuyó a la idea de evaluación dirigida por teoría y evaluación comprensiva —*theory-driven evaluation & comprehensive evaluation*— descrita

⁴ La evaluación tuvo un desarrollo importante en el campo de la educación, por lo que muchos estudios provienen del ámbito educativo.

como la sistemática aplicación de procedimientos de la investigación social en estimar la conceptualización y diseño, la implementación y la utilidad de los programas de intervención social. Esto envuelve la construcción de una detallada teoría del programa que es usada como guía de la evaluación. Rossi también sugirió que las evaluaciones se modelaran de acuerdo con los programas o políticas, *fitting evaluations to programs*.

Huey-Tsyh Chen (1987) desarrolló el concepto y la práctica de la *evaluación dirigida por la teoría*. Cuando los experimentos controlados no muestran efectos no se puede saber si se debe a relaciones de causalidad escasamente construidas, insuficientes niveles de tratamiento o pobre implementación. Por lo tanto, es importante realizar los experimentos y los cuasi experimentos con conocimiento y teoría previos a fin de construir modelos de procesos de tratamiento e implementación, que produzcan evaluaciones que sean más eficientes y que provean mayor información sobre cómo lograr los efectos deseados. Se trata de construir modelos defendibles y plausibles de cómo puede esperarse que los programas/políticas trabajen. Esto nos permitirá encontrar efectos secundarios y consecuencias no esperadas.

Para Carol Weiss (1991) la evaluación es un sinónimo de investigación cuidadosa, pero reconoce que es una actividad política que afecta y es afectada por el contexto político, debido a que los programas y planes son creados y mantenidos por fuerzas políticas; las decisiones se toman a nivel político y la evaluación misma tiene connotaciones políticas.

b) Valoración

La segunda gran rama teórica está relacionada con la función de valorar, de hacer juicios que tiene el evaluador. Michel Scriven (1986) sostiene que las sociedades necesitan valorar y que ésta es una función del evaluador; no sólo se trata de que la evaluación proporcione información. Este autor está en desacuerdo en tomar los objetivos del programa o política como punto de partida de la evaluación; no hacerlo y darle libertad al evaluador ayudará a identificar los logros reales del programa.

Otro teórico, Elliot Eisner (1998), desde una óptica diferente también le da importancia a la valoración. La evaluación implica realizar *juicios de valor* de la cualidad de un objeto, proceso o situación, y como los instrumentos cuantitativos no son suficientes, es necesario entender el significado de lo que es *conocido*. Sus dos conceptos, *connoisseurship* y *criticism*, se relacionan con la capacidad del experto de hacer crítica durante la evaluación, incluidas en ésta la descripción, la foto de la situación o el proyecto; la expectación en la que el evaluador trata de entender lo que vio, a partir de su conocimiento de experto y, por último, la percepción, en la que el evaluador se preguntará ¿cuál es el valor de lo que sucede? Para Eisner la esencia de este proceso es la selectividad; el *connoisseur* no describe todo, selecciona a partir de juicios de valor.

Otro enfoque de la valoración aplicada a la evaluación es la *evaluación por adversarios*, *adversary evaluation*, en la que se emplean evaluadores para representar dos puntos de vista opuestos sobre los datos con dos interpretaciones finales. Por otra parte Robert Stake (2001) considera que ningún elemento tiene verdadero valor; el conocimiento está sujeto al contexto; la perspectiva de los *stakeholders* es parte integral de la evaluación y los estudios de caso son el mejor método de representar creencias y valores o de reportar los resultados de las evaluaciones.

Asimismo, las perspectivas tienen que representarse en la evaluación, pero el contenido del reporte de la valoración lo hace el evaluador a discreción.

El evaluador debe reconocer la multiplicidad de perspectivas de los interesados en un programa o política determinados, y presentarlos y tomar en cuenta los diferentes contextos político-burocráticos, autocráticos o democráticos en que se desarrollan las evaluaciones (MacDonald, 1979).

Ernest House (1991, 1993) arguye que la evaluación nunca es neutral; se debe inclinar por la justicia, la equidad y lo correcto, al considerar de manera prioritaria los intereses y las necesidades de los que tienen menos poder.

Otros teóricos —E. Guba e Y. Lincoln (1989)— ven a los *stakeholders* como los responsables de valorar, mientras que el papel del evaluador es el de facilitar las negociaciones entre individuos que reflejan diferentes realidades. Los evaluadores son vistos como orquestadores que buscan el consenso en el proceso de negociación.

En síntesis, diversos teóricos enfatizan la función de valoración, de hacer juicios en los procesos de evaluación, aunque los evaluadores lleguen a ella por medio de diferentes maneras y responsabilidades.

c) Uso

Los teóricos de *evaluación orientada por decisiones* sostienen que las evaluaciones se deben diseñar para asistir a los *stakeholders* en la toma de decisiones. Daniel Stufflebeam (2007) desarrolló en este contexto el modelo CIIP, así llamado por las cuatro tipos de evaluación, que incluye: contexto, insumos, proceso y producto. La evaluación de contexto identifica las necesidades para decidir; la de insumo examina las decisiones sobre estrategias y diseños; la de proceso reconoce las deficiencias para mejorar la implementación, y la de producto mide los resultados para la toma de decisiones respecto de la continuación o rediseño del programa o política.

Stufflebeam afirma que para mejorar los servicios de un proyecto o programa la evaluación tiene que ayudar a los responsables de las políticas a asignar recursos a los programas que mejor sirvan a los clientes. Los evaluadores deberían conducir su trabajo de acuerdo con las normas siguientes: utilidad, factibilidad, propiedad, exactitud. Los contratos para una evaluación deben detallar qué se debe hacer, cómo, por quién, y cuándo, para que las partes estén obligadas a adherirse a dichas condiciones.

Para apoyar a los administradores de programas, Provus (1971) considera que la evaluación, como proceso continuo, debe desarrollar cuatro etapas: a) *definición* de metas, procesos y recursos; b) *instalación*, identificación de las discrepancias en la implementación del programa; c) *proceso*, para determinar el grado de logro de los resultados a corto plazo, y d) *producto*, el grado en que se han alcanzado los objetivos de largo plazo. Una quinta etapa opcional podría ser el análisis de costo beneficio.

El interés en evaluaciones de grandes programas condujo a concentrar la atención en el uso que administradores y *policy-makers* podrían dar a las evaluaciones para mejorar la administración. Al respecto, Joseph Wholey sostiene que obtener información a partir de evaluaciones es costoso, por lo que propone un procedimiento de cuatro etapas para la “compra secuencial de información”. La primera etapa es: examen de evaluación que consiste en un examen o preexamen inicial de la factibilidad de realizar la evaluación en términos de organización y de los tópicos a ser estudiados. Además, considerar si los resultados de la evaluación serán útiles para que los administradores mejoren sus programas.

Las otras tres etapas del proceso de Wholey son: a) retroalimentación rápida de la evaluación, que se enfoca a examinar la existencia y la facilidad de recolectar la información; b) monitoreo del desempeño, que mide el desempeño del programa/política, en comparación con el desempeño esperado o previo, y c) evaluación intensiva, que utiliza comparación o grupos de control para estimar mejor la efectividad de las actividades del programa en producir los efectos observados. Wholey enfatiza los beneficios potenciales —la utilización instrumental y en el corto plazo— de la evaluación para la administración.

Una visión más amplia de los usos de la evaluación es la de Michael Patton, para quien la evaluación no debe estar únicamente relacionada con quienes deciden qué políticas implementar, sino que el evaluador tiene la obligación de tratar de asegurar su utilización. El evaluador debe buscar usuarios reales de la evaluación, *potenciales usuarios primarios*.

Patton (1997) propuso un mayor desarrollo y énfasis en el papel del evaluador en el uso de la evaluación. Presentó un proceso de *evaluación focalizada en la utilización* en cinco pasos: a) identificación de los potenciales usuarios; b) compromiso de los potenciales usuarios en la evaluación y en la utilización de la evaluación; c) participación en métodos, diseño, y medición; d) compromiso activo y directo en interpretar los resultados y hacer juicios; y e) tomar decisiones sobre disseminación adicional.

Para Patton (1997) se trata de identificar personas que tengan interés en el programa/política a evaluar. El evaluador debe ser especialmente activo en desarrollar el compromiso de los potenciales usuarios, al comprometerlos en todas las etapas de la evaluación. El evaluador debe ser: a) activo en identificar potenciales usuarios; b) reactivo en aprender continuamente de la situación evaluada, y c) adaptable, en modificar las preguntas y diseños de la evaluación a la luz del mayor entendimiento de la situación y de las condiciones cambiantes.

A fines de los años noventa las teorías sobre uso de las evaluaciones aumentaron; diversos teóricos trabajaron sobre los modelos de evaluación participativa. J Bradley Cousins sostiene que la utilización de ésta debe lograrse al comparar con los posibles usuarios *to get buy in*, para lo que es necesario lograr que el personal del programa participe en la evaluación. La evaluación participativa práctica se define como la investigación social aplicada que considera al personal capacitado en la toma de las decisiones, basada en la práctica en la colaboración. Las responsabilidades son compartidas entre el evaluador y el personal del programa/política.

Para Jean King la evaluación es un proceso de investigación sistemática que provee información confiable acerca de las características, actividades y resultados de un programa o política para un propósito útil. Para que una evaluación participativa sea exitosa debe haber una

estructura de poder aceptable; compartir las explicaciones de las experiencias entre los participantes; haber voluntarios y líderes; tiempo suficiente y elevado grado de confianza interpersonal y organizacional. La construcción de confianza es fundamental para el éxito de la evaluación participativa.

David Fetterman trasciende y define *Empowerment Evaluation* como el proceso de estimular la determinación propia entre los que participan en la evaluación de un programa. Esto incluye capacitación, facilitación, apoyo, ilustración y liberación. El evaluador es un entrenador y facilitador que capacita a los participantes para que evalúen sus propios programas y la asesoría en el diseño de la evaluación es una forma efectiva de empoderamiento. La evaluación como empoderamiento es un proceso continuo que permite el cambio de roles y expectativas tradicionales.

3. Evolución reciente de los trabajos sobre evaluación de políticas

a) Evolución reciente

En los últimos siete años se ha presentado una producción considerable de estudios, manuales, libros, folletos y otros trabajos acerca del tema. En estos trabajos se incluyen diversos estudios académicos, en buena parte productos de organismos bilaterales de Japón, Alemania, Estado Unidos, Reino Unido, Suecia (cuyo manual de 2004 se destaca por sus alcances), y multilaterales: Banco Mundial y Naciones Unidas (PNUD, CEPAL, FAO, FIDA). Como respuesta a ellos se ha generado una gran cantidad de evaluaciones de proyectos, programas y políticas, en muchos casos financiados por estos organismos, con énfasis en acciones de combate a la pobreza (véase el anexo 1 con referencias de diversos organismos).

Por su parte, varias asociaciones y organizaciones privadas están dedicadas al desarrollo de metodologías, criterios y prácticas de evaluación; por ejemplo, The American Evaluation Association, The Kellogg Foundation y Monitoring and Evaluation News, entre otras.

La mayoría de los estudios y guías mencionados incluyen metodologías y manuales que se refieren a temas en general, o del ámbito educativo, y no a políticas rurales o agropecuarias. Hay un énfasis en los temas de pobreza, empoderamiento, género y metas del milenio.

Actualmente la práctica de las evaluaciones redundante en su utilidad. Su propósito es conocer cómo las acciones públicas y privadas pueden usarse para fortalecer la efectividad de la política, en nuestro caso de la política rural. Así, la evaluación se considera como un insumo para la reforma de políticas; se trata de que contribuya a identificar las restricciones institucionales para superarlas, los *espacios vacíos* para proponer mecanismos e instancias adecuadas.

Por otra parte, en la actualidad, “el desafío para la evaluación es desenmascarar los intereses de los buscadores de rentas e informar al público acerca de las opciones para corregir las fallas del mercado” (Chelimsky y Shadish, 1997). La selectividad es también otro rasgo de las evaluaciones. Por una parte, conviene realizar evaluaciones estratégicas de las principales

políticas y de estructuras institucionales seleccionadas; por otra parte, se debe ser selectivo en los temas incluidos en cada evaluación.

Así, en el caso de la evaluación de políticas, el desafío de la evaluación es identificar los errores o las omisiones de la agenda de la reforma —rural— de los países en desarrollo, ya sea una falla institucional o del gobierno y asesorar con base en los resultados. En el caso de la evaluación de programas, el reto es establecer la causalidad de los efectos de los programas o instrumentos de la política, así como medir dichos efectos. También la evaluación comprende el ámbito de la implementación (procesos) de los programas.

b) El énfasis en evaluación de resultados

“One of the ways from which you can distinguish different evaluation approaches is by what they take as their bottom line for the evaluation. For me, it is the pragmatic use of evaluation findings and the evaluation process. In other words, it is that **the evaluation is designed and implemented in a way that really makes a difference to improving programs and improving decisions about programs**. So the bottom line in my approach is *use* — that’s the reason why my approach is called utilization focused evaluation. Michael Quinn Patton.

Como se comentó en la parte teórica, en los años 2000 se ha enfatizado el uso de las evaluaciones y, en especial, a la orientada por resultados. El manual del SIDA, desde su título, acentúa este concepto *Looking Back, Moving Forward* (SIDA, 2007), al igual que el manual de OECD/DAC, *Evaluations and Results-Based Management*.

La evaluación es vista como la forma de aprender de la experiencia anterior para mejorar la prestación de servicios, planificar y asignar los recursos de manera más eficiente y demostrar los resultados a la sociedad y a los cooperantes. La rendición de cuentas, que ha sido un fundamento de las teorías de la evaluación, es ahora uno de los propósitos importantes de estos procesos. La evaluación es pues un mecanismo útil para funcionarios públicos, responsables de programas rurales y la sociedad civil.

En el contexto de la evaluación orientada por resultados, dos modalidades de evaluación de la política pública merecen consideración particular. En primer lugar, el tema de las encuestas de servicios al cliente para evaluar el desempeño de los servicios gubernamentales (identificación de obstáculos, opiniones sobre calidad y suficiencia y la capacidad de respuesta de los funcionarios). En segundo lugar, los estudios de seguimiento del gasto público rural para determinar hasta qué punto los recursos llegan a los destinatarios, identificar problemas existentes y ofrecer pruebas de retrasos y *filtraciones*. Estos estudios ayudan a mejorar la gestión y favorecen la rendición de cuentas (Banco Mundial, 2004).

Hay que recordar que, como en otros campos, no hay un único y “mejor modelo” de sistema de evaluación. Por el contrario, las evaluaciones tienen que ser “hechas a la medida” de las necesidades, dependen del propósito y del contexto en que se realizan, y lo importante es evaluar el impacto del gasto anual de los programas, y no sólo de proyectos, así como de los programas estratégicos de las instituciones que muestran efectos sobre otros programas y proyectos.

También conviene comentar la importancia que ahora se pone en la asociación de las actividades de monitoreo y de evaluación en diversos estudios para asegurar los efectos y objetivos.⁵

El desarrollo de la institucionalidad debe soportar a la evaluación orientada por resultados. Así, el mecanismo institucional encargado de diseñar las evaluaciones, implementarlas en varios casos y de velar para que se lleven a cabo sus recomendaciones es de suma importancia. Dicho mecanismo es responsable de facilitar el diseño de las políticas y procedimientos de evaluación y seguimiento, definir el alcance de la evaluación, organizar el equipo evaluador y proporcionar y ajustar a cada evaluación las metodologías a utilizar de acuerdo con cada contexto. También debe velar porque las unidades evaluadas respondan a las observaciones y recomendaciones de la evaluación y de dar seguimiento a la implementación de las recomendaciones.

4. Enfoques institucionales de la evaluación de políticas y proyectos

a) Banco Mundial

El Banco Mundial tiene una valiosa historia de más de 30 años de evaluación de sus operaciones en pro del desarrollo. Inició en 1970 y en 1975 la función de evaluación se independizó y formó el Grupo de Evaluación Independiente, antes Departamento de Evaluación de Operaciones del Banco Mundial. El Grupo cuenta con una larga experiencia que ha crecido con el aumento, diversificación y cambios de prioridades del Banco. Así, ha realizando evaluaciones a diversos proyectos sobre infraestructura, agricultura, medio ambiente, salud, educación y pobreza, que se presentan en su sitio web (Picciotto, 1997).

En el sitio del Grupo de Evaluación Independiente (GEI) hay diversas notas metodológicas e instrumentos sobre monitoreo y evaluación, acerca de la creación de sistemas institucionales para realizar estas actividades, y se presentan varias experiencias de evaluación de proyectos.

De la experiencia de evaluaciones del desarrollo del GEI se pueden extraer dos conclusiones importantes: primero, con relación al contexto. Las evaluaciones del desarrollo ocurren en circunstancias difíciles, con datos escasos, de no tan buena calidad y con información de base escasa. Generalmente, la capacidad institucional local para realizar evaluaciones es débil, como también lo es la capacidad de aplicar los resultados de las evaluaciones. Las diferencias culturales pueden ser un obstáculo. Sin embargo, para compensar esto hay un gran interés por lecciones relevantes de la experiencia (Ingram, 2003).

La segunda conclusión se refiere al concepto cambiante del desarrollo. Se han presentado modificaciones importantes en los temas prioritarios para el desarrollo a lo largo del tiempo. Las evaluaciones han ampliado el concepto de capital productivo, al capital humano y al capital

⁵ Estas ideas se han resumido de Banco Mundial (2004).

social; la estabilidad macroeconómica y marcos de política apropiados son considerados necesarios para el crecimiento, y la reducción de la pobreza se ha convertido en el mayor objetivo, con nuevos instrumentos en educación, salud y redes de protección social. El buen gobierno y la participación ciudadana son elementos centrales de las estrategias, que se focalizan en países más que en proyectos o sectores; y *partnerships*, las asociaciones público-privadas son importantes modalidades de hacer negocios. Por último, nuestras estrategias están concentradas ahora en **resultados**, incluidos en las Metas del Milenio (Ingram, 2003).

Desde los años noventa el Banco Mundial ha reorientado sus actividades para privilegiar la cooperación a los países en el combate a la pobreza. El desafío aún es considerable. Las evaluaciones han desempeñado un papel importante en este esfuerzo a fin de aumentar su efectividad, elevar el número de programas nacionales con logros satisfactorios y reducir el costo de sus programas (Banco Mundial, 2005). En este contexto los métodos y herramientas de la evaluación de impacto han sido desarrollados por el GEI.

En 2007 el Grupo publicó el *Sourcebook for Evaluating Global and Regional Partnership Programs. Indicative Principles and Standards*, con el propósito de mejorar la independencia y calidad de las evaluaciones y para ampliar la relevancia y efectividad de los programas a partir de los principios indicativos y criterios contenidos en el trabajo. El mismo está dirigido a gobiernos, evaluadores y tomadores de decisiones (Grupo Independiente de Evaluación, 2007).

b) La FAO

El Servicio de Evaluación de la FAO inició sus operaciones en 1968 y es parte de la Oficina de Programación, Presupuesto y Evaluación. Entre sus funciones se encuentran: el desarrollo del sistema de evaluación; evaluación de información técnica y analítica y asesorías; asegurar la evaluación independiente de proyectos en terreno; mantener un banco de datos con los resultados de las evaluaciones de proyectos; promover y dar retroalimentación sobre evaluación para el aprendizaje organizacional y el mejoramiento de la institución y, de especial importancia, proveer recursos técnicos para los países miembros en programación y monitoreo y evaluación de proyectos.

Este organismo cuenta con una amplia experiencia de evaluación, autoevaluación y evaluación independiente de sus proyectos, políticas y programas. Ha publicado diversas notas metodológicas y guías acerca del tema. Su Manual de autoevaluación AE incluye entre otros: un marco conceptual, procedimientos sobre alcances, roles y secuencia de la AE, tipos, criterios de calidad y exposición de diversas técnicas. Contiene orientaciones para realizar evaluaciones de mayor alcance de la FAO misma, así como una gran cantidad de evaluaciones de sus proyectos de cultivos y productos pecuarios.

La FAO lleva a cabo cinco tipos de evaluaciones: las principales de los programas individuales de la institución; las de gran alcance, del trabajo de FAO a nivel de cada país; las de proyectos extrapresupuestarios, incluidos los trabajos de emergencias; las de programas y proyectos extrapresupuestarios particulares, y las autoevaluaciones.

En 2003 la FAO publicó unas *Orientaciones para la Autoevaluación* para ser aplicadas por la institución en la revisión de sus actividades en períodos de alrededor de seis años y con una perspectiva amplia de objetivos y resultados. Esto con el propósito de mejorar sus programas, obtener lecciones de los proyectos y rendir cuentas a los *stakeholders*.

El caso de la FAO es un claro indicador de la relevancia de la evaluación para las instituciones internacionales de cooperación, ya que en 2007 dio a conocer una amplia evaluación independiente externa realizada a la organización (FAO, 2007a). En la evaluación se concluyó que era necesario que la institución cambiara para hacer frente a los desafíos del siglo XXI y revisara su plan de acción en cuatro áreas:

- i) Crear un nuevo marco estratégico.
- ii) Invertir en gobernabilidad.
- iii) Cambiar la cultura institucional, y
- iv) Reformar los sistemas administrativos y gerenciales, y reestructurar en busca de eficiencia y efectividad.

En 2007 la FAO publicó el *Programme Evaluation Report*, que contiene información acerca de los arreglos, políticas y métodos de la Institución; se coordinó con el sistema de Naciones Unidas para la evaluación e incluye las evaluaciones realizadas y planeadas en el bienio (FAO, 2007a).

Recientemente la FAO ha dirigido sus trabajos de evaluación a los temas de arreglos institucionales, políticas y métodos, evaluaciones relacionadas con respuestas de emergencia y rehabilitación; y responsabilidades y procedimientos de respuesta y seguimiento a evaluaciones (FAO, 2007b).

En el caso de México, la FAO, conjuntamente con SAGARPA, ha desarrollado un trabajo significativo de evaluación de políticas rurales a lo largo de la década pasada. En el siguiente capítulo se presenta una síntesis analítica de las metodologías, valoraciones y usos de las evaluaciones realizadas en el período 1996-2007.

c) **El IICA**

El Instituto Interamericano de Cooperación para la Agricultura (IICA) ha trabajado en el desarrollo de enfoques y metodologías de evaluación de políticas públicas, con el propósito de apoyar a los gobiernos de la región para contar con instrumentos de medición que permitan hacer un seguimiento a sus estrategias e inversiones. El IICA ha aplicado específicamente enfoques innovadores de evaluación en el desarrollo del Sistema de Monitoreo y Evaluación del Desarrollo Rural Sustentable Municipal en México y en el Sistema de Gestión Estratégica para la política de desarrollo territorial rural en Brasil.

El Instituto ha realizado modelos y metodologías para los procesos de evaluación y seguimiento de los impactos territoriales de las políticas públicas rurales, en particular para examinar los impactos en el desarrollo rural municipal. En el tercer capítulo de este estudio se presentan los trabajos del IICA en el Sistema de Monitoreo y Evaluación del Desarrollo Rural Municipal de México.

d) La CEPAL

Desde la década de los sesenta, esta institución y su organismo de planificación, el Instituto Latinoamericano y del Caribe de Planificación Económica y Social (ILPES), han desarrollado varios trabajos en evaluación de proyectos, programas y políticas y han hecho una extensa labor en asesoría y capacitación en materia de evaluación a los países latinoamericanos. En los últimos años, han escrito algunos manuales sobre evaluación de impacto, cuyo propósito es exponer los objetivos, los principales conceptos involucrados y métodos de estimación utilizados, así como presentar evaluaciones selectivas de programas que han sido valorados con estas metodologías, incluidos los programas centrados en mitigar la pobreza estructural, programas de empleo diseñados para tiempos de crisis y programas de capacitación de jóvenes en América Latina (CEPAL, 2005).

El ILPES ha examinado la función de evaluación de planes, programas, estrategias, y proyectos y ha concluido que “la necesidad de evaluar las acciones del Estado en los países de América Latina surge de tres hechos:

i) La importancia de introducir mayor racionalidad a los ajustes o correcciones fiscales, que suponen un adecuado conocimiento de las actividades del sector público y de su impacto para efectuar los recortes de manera diferenciada y según prioridades bien establecidas.

ii) El alto costo político que pagan los gobiernos cuando la opinión pública advierte escasa eficiencia en los programas públicos.

iii) La mayor demanda de transparencia y de responsabilidad por parte de sociedades civiles más democráticas, que debería generar una mayor demanda de evaluación, pero las restricciones institucionales la dificultan. La pregunta básica sería cómo aumentar la demanda de evaluación; dentro del necesario ejercicio de evaluación de la gestión pública se pueden distinguir dos tipos de evaluación: el referido a las organizaciones públicas y el otro cuyo objeto es más amplio, que serían los marcos macroinstitucionales en los que dichas organizaciones se desenvuelven. La evaluación de las organizaciones públicas puede realizarse bajo dos formas: la autoevaluación y la evaluación estratégica (Weisner, s/f).

En agosto de 2007 el ILPES publicó un trabajo sobre evaluación de programas (ILPES, 2007) en el que se muestran aspectos específicos de los procesos de evaluación ente otros tópicos conceptuales, metodológicos, de institucionalidad y de diseño operativo, relevantes con respecto a su fin en el ámbito de las políticas públicas. Algunos temas revisados por el estudio son: los resultados esperados e identificación de hipótesis a comprobar; mediciones de desempeño; población beneficiaria; aspectos cualitativos; herramientas de evaluación; disponibilidad de información y limitaciones del proceso de evaluación; incorporación del enfoque o la perspectiva

de género en los programas públicos y en su evaluación; análisis de resultados y compromisos, y aspectos de institucionalidad y gestión.

5. Resumen

En este capítulo se ha resaltado la importancia de la evaluación de programas, proyectos y políticas, en particular rurales, para mejorar su eficiencia, efectividad, la rendición de cuentas, la asignación de recursos, y en especial la política pública. La evaluación contribuye a formular estrategias y políticas rurales mejor diseñadas y de mayor impacto; además, facilita una apreciación mejor de los logros de la cooperación internacional. Temas prioritarios de las agendas nacionales e internacionales, como la reducción de la pobreza y las tendencias de los precios internacionales y la oferta de los alimentos, se abordan de manera más efectiva a partir de los resultados de las evaluaciones.

Se han revisado de manera somera los fundamentos teóricos del proceso de evaluación y sus ramas principales de desarrollo, los métodos, el uso de la evaluación y la valoración que hace el evaluador o los interesados en el proceso evaluado. Ha quedado en evidencia la gran cantidad de investigaciones y desarrollos teóricos basados en la investigación social científica y la rendición de cuentas y control.

De manera muy sintética se han esbozado algunos desarrollos del pensamiento sobre evaluación. En el caso de los métodos, la evaluación orientada por objetivos, el diseño experimental y cuasi experimental, la investigación evaluativa, la consideración especial por el contexto y los *stakeholders*, la evaluación dirigida por teoría y las diseñadas para cada programa.

En cuanto a la rama teórica de valoración, la capacidad que tiene el evaluador como experto para juzgar al programa, política o proyecto; la evaluación por adversarios que complementen el juicio, la consideración de múltiples contextos y las perspectivas de los interesados en la evaluación. El hecho de que la evaluación no es neutral y tiene que atender las necesidades y los intereses de los que tienen menos poder.

El uso de los resultados de la evaluación ha sido pensado para la toma de decisiones, para los *stakeholders* y para los administradores de programas. Los evaluadores tienen que *vender* la evaluación entre estos usuarios e identificar nuevos.

Para finalizar, se ha resumido el aporte valioso que han hecho el Banco Mundial, la FAO, la CEPAL y el IICA en cuanto a metodologías, estándares, realización de evaluaciones y lecciones aprendidas durante los últimos 40 años.

En una época de crisis global, de astringencia de recursos y de gran volatilidad de mercados, entre ellos los agropecuarios, la evaluación de políticas y proyectos públicos rurales puede contribuir significativamente a ser más eficientes en el uso de los fertilizantes, el agua, la energía y la tierra, como afirma Alexander Evans.

En el siguiente capítulo se examinan las valoraciones realizadas al principal Programa de fomento productivo rural de México durante los últimos 10 años. Se comentan en especial las metodologías empleadas, sus logros, limitaciones y algunas lecciones aprendidas en los procesos de evaluación.

II. APRENDIZAJES DE LA EVALUACIÓN DE PROGRAMAS DE FOMENTO: EL CASO DE ALIANZA PARA EL CAMPO

1. Introducción

Como se mencionó en el capítulo anterior, la evaluación de programas públicos tiene como uno de los objetivos prioritarios proveer información acerca de la gestión y el desempeño de las intervenciones con objeto de retroalimentar a los responsables en la toma de decisiones sobre la eficiencia y eficacia en la ejecución de dichos programas. Con el fin de compartir los aprendizajes derivados de la experiencia de evaluación de Alianza para el Campo, en este apartado se presenta, de manera sintética, la evolución de diversos aspectos metodológicos de la evaluación de este programa.

2. La Alianza para el Campo

Alianza para el Campo (APC) fue el principal programa de fomento productivo rural de la política sectorial del gobierno federal mexicano en las dos administraciones federales pasadas. Vigente de 1996 a 2007, en Alianza se indujo una inversión total de 131.000 millones de pesos ⁶ durante ese período; la tasa media anual de crecimiento de los recursos del Programa fue de 4,9%.

En su diseño, Alianza se distinguió por su carácter descentralizado, su orientación a la demanda por parte de los productores con respecto al tipo de inversiones, así como por el esquema de concurrencia de recursos de los gobiernos federal y estatal con los de los propios beneficiarios. Con relación a este último aspecto, en los 12 años de Alianza, el promedio de la participación federal en los recursos fue de 44%, la de los estados de 16% y la de los beneficiarios de 40%.

Alianza entregó subsidios a fondo perdido para la formación de capital físico, humano y social por medio de siete programas: a) fomento agrícola; b) fomento ganadero; c) acuacultura y pesca, d) desarrollo rural; e) sanidad e inocuidad agroalimentaria; f) Sistema Nacional de Información para el Desarrollo Rural Sustentable, y g) Fondo de Estabilización, Fortalecimiento y Reordenación de la Cafecultura.

Del total de los recursos de Alianza, 76% se destinó a bienes privados (adquisición de bienes de capital), de los que 95% se canalizó a la producción primaria. El 24% restante de la inversión se destinó a bienes públicos (investigación y transferencia de tecnología, sanidad vegetal y animal, inocuidad de alimentos, capacitación a productores e información de mercados). Los 500.000 beneficiarios anuales, en promedio, recibieron de 1996 a 2006, a través de Alianza, principalmente tractores (47.316), sementales (170.000), vaquillas (450.000) y, entre

⁶ A precios de 2006.

otros apoyos entregados, se atendieron 15 millones de hectáreas de agostadero y 272,757 con riego tecnificado.

Alianza buscó incidir en la integración de cadenas productivas la reconversión productiva y la atención a grupos y regiones prioritarias. En ese sentido, Alianza contribuyó de manera importante a impulsar la creación de los Consejos de Desarrollo Rural Sustentable, así como la integración de cadenas agroalimentarias y pesqueras mediante la integración de los Comités Sistema Producto y la inversión en posproducción.

Los impactos generados por las inversiones de Alianza giran principalmente alrededor del efecto en el ingreso de los productores, propiciado por los cambios en la escala de producción y, en menor medida, por mejoras en la productividad, debidos a su vez al incremento en la capitalización y al cambio tecnológico.

3. La evaluación de Alianza para el Campo

Dado su alcance y la importancia presupuestal de Alianza, el Gobierno de México solicitó, en 1999, a la Organización de las Naciones Unidas para la Agricultura y la Alimentación (FAO) la evaluación del conjunto de programas de Alianza para el Campo correspondiente al ejercicio de 1998, que fue el tercer año de operación de Alianza. En un principio la evaluación se enfocó a estimar los resultados e impactos de los diferentes programas que la comprenden y fue a partir de 2001 que se incorporó la evaluación de procesos o gestión de dichos programas.

a) Evaluación sumativa (impactos)

La evaluación de Alianza representó varios retos, ya que el Programa no contemplaba, desde el principio, a la evaluación como parte integral de su diseño. Ello implicó que Alianza no contara con indicadores de desempeño definidos *ex ante*, tanto para la gestión como para los resultados e impactos. En ese sentido, tampoco se dispuso de información contrafactual, a partir de líneas de base, que sirviera para valorar los impactos, y que fuera posible aislar, de manera reflexiva, los efectos de Alianza sobre el ingreso, el empleo, la productividad, la capitalización y otras variables sobre las que se midieron los impactos del Programa; o bien a partir de grupos de comparación que permitieran, mediante pareo entre beneficiarios y no beneficiarios, establecer la causalidad de los impactos de Alianza. Este factor, impuesto por el contexto, constituyó una restricción fundamental en cuanto a las opciones metodológicas de evaluación.

Otro reto que presentó la evaluación de Alianza, en términos metodológicos, fue el carácter múltiple de tratamientos de este Programa, que otorgaba una amplia gama de tipos de apoyo, tanto de bienes públicos como privados, ya fuera en el ámbito agrícola, pecuario, pesquero y acuícola, como en el no agropecuario. De hecho, hasta antes de 2003, Alianza cambiaba cada año, tanto en su diseño como en su normatividad, y modificaba también los conceptos de apoyo otorgados.

Además de los elementos cambiantes de diseño antes mencionados, la población atendida era diferente en cada ejercicio del Programa. Ello generaba dificultades metodológicas

adicionales, pues de haberse contado con una línea de base, sólo habría sido posible evaluar los impactos de la población atendida en el período base y, de haberse optado por grupos de comparación, éstos se habrían agotado a medida que avanzaba la cobertura de Alianza.⁷

Por otra parte, la normatividad mexicana aplicable a Alianza establecía que las evaluaciones de programas federales debían realizarse anualmente y cuando 60% de los recursos ya hubiesen sido ejercidos. Sin embargo, este límite no siempre coincidió con el tiempo requerido para la maduración de los impactos debidos a las inversiones, ni con los ritmos de operación de la Alianza. Aunado a ello, al momento de la evaluación, las inversiones otorgadas a los beneficiarios tenían cuando mucho siete u ocho meses de haber sido recibidas, y aumentaron así las limitaciones para determinar impactos, sobre todo en aquellas inversiones que requieren de una mayor maduración.

Además de las consideraciones anteriores, cabe señalar que Alianza para el Campo fue un programa de operación descentralizada. Es decir, diseñado y normado por el gobierno federal, pero operado en coparticipación con los gobiernos de los estados; cada estado contaba con la facultad para focalizar la población objetivo y priorizar el tipo de apoyos otorgados de acuerdo con la problemática específica de cada Estado. Ello implicó que en cada Estado se llevara una evaluación local de los programas de Alianza operados estatalmente.

En ese contexto de retos y oportunidades, la evaluación de impactos de la Alianza para el Campo que realizó el Proyecto FAO-SAGARPA consistió en una evaluación cuantitativa basada en el diseño de indicadores a partir de datos recopilados mediante encuestas de percepciones a beneficiarios del Programa. Dicha evaluación fue aplicada a una muestra representativa a nivel estatal de cada uno de los siguientes programas de Alianza: Fomento Agrícola, Fomento Ganadero y Desarrollo Rural. La evaluación nacional de los programas de Alianza, por su parte, se llevó a cabo a través de la agregación de los tamaños estatales de muestra, lo que aseguró la representatividad a nivel nacional.

Debido a que el diseño de Alianza cambiaba anualmente, hasta antes de 2003, el cuestionario a beneficiarios se ajustó igualmente para dar acomodo a los indicadores de impacto acordes con los conceptos de apoyo contemplados, y a los objetivos establecidos, en cada ejercicio del Programa. De esa manera, el cuestionario a beneficiarios evolucionó a lo largo del tiempo en función de los temas de evaluación definidos cada año, que a su vez estaban determinados por los cambios anuales de Alianza.

Cabe señalar que en los primeros años, parte importante de la evaluación consistió en conocer la opinión de los beneficiarios acerca de la calidad del servicio brindado por los funcionarios, la calidad y oportunidad de los apoyos recibidos, las necesidades de servicios complementarios, tales como asistencia técnica y el nivel de conocimiento del Programa. Sin embargo, los indicadores relativos a esos temas mostraron un desempeño más alto que el observado en la realidad debido principalmente a que los beneficiarios encuestados advirtieron que sus posibilidades de acceso a posteriores apoyos podrían verse afectadas por la manera de responder el cuestionario, razón por la que paulatinamente se eliminaron dichos temas, se acotó

⁷ Ésas eran, al menos, las razones que en ese momento se tenían como limitantes para la evaluación rigurosa de impactos.

el ámbito de evaluación a la medición de impactos y se focalizó la obtención de información sobre variables ligadas directamente a los principales objetivos del Programa.

Asimismo, con respecto a los temas de evaluación vinculados con los objetivos de los programas de Alianza, en una primera etapa las preguntas se dirigieron a obtener información a partir de las opiniones y percepciones de los beneficiarios y los impactos se midieron con base en ellas. Así, por ejemplo, para la medición del impacto en el ingreso, se preguntó a los beneficiarios su percepción acerca de los cambios en la producción y el rendimiento. El indicador de ingreso medía el porcentaje de beneficiarios que había advertido cambios en esas variables, sin poder aislar el efecto de Alianza de otros factores que podrían haber incidido en ellos ni cuantificar su dimensión. Con ello, la objetividad y la precisión de los indicadores resultó limitada.

Se buscó incrementar la calidad y objetividad de la información obtenida mediante la encuesta a beneficiarios, por lo que se incluyeron preguntas sobre la situación de variables en distintos momentos del tiempo a fin de hacer una estimación cuantitativa de su comportamiento. Por ejemplo, en el caso de la medición del ingreso se obtenía información acerca de las variables de producción y precios antes y después de la recepción del apoyo. A su vez, los cambios en producción se midieron a través de variaciones en las escalas de producción y en los rendimientos, con lo que fue posible desglosar y cuantificar la contribución de cada una de estas variables a los cambios en el ingreso. De esa manera, ya no se reportaron percepciones sobre los cambios en determinadas variables, sino que se incorporaron preguntas para medir la situación antes y después del apoyo recibido, además de que se identificaron y aislaron de mejor manera los impactos de Alianza de los efectos de otras variables. En ese sentido, el instrumento de colecta de información en campo cambió gradualmente, focalizó la información relevante y precisó una orientación hacia un conjunto de temas e indicadores básicos, que fueron:

- Identificación del beneficiario
- Perfil del beneficiario
- Características de los apoyos recibidos
- Ingreso y empleo en las actividades productivas en las que se utilizó el apoyo
- Cambio tecnológico en las actividades productivas en que se utilizó el apoyo
- Capitalización de la unidad productiva
- Producción y productividad en la actividad en la que se aplicó el apoyo
- Desarrollo de capacidades y de organizaciones económicas

Por otra parte, y dentro de una iniciativa del Proyecto de Evaluación por aislar los efectos de las intervenciones, a fin de contar con mayor rigor en la evaluación de impactos, se realizó, con un grupo piloto, un ejercicio de evaluación bajo una metodología cuasi-experimental. La existencia de tratamientos múltiples y la falta de información ⁸ acerca de la población no beneficiaria fueron, en esencia, las dificultades encontradas en la aplicación de esta metodología. La evaluación piloto fue diseñada inicialmente para aplicar el método cuasiexperimental en dos programas contrastantes: Desarrollo Rural y Tecnificación del Riego, ⁹ en dos estados también

⁸ Información acerca de variables observables relevantes para el pareo con beneficiarios, pero sobre todo la ausencia de información para ubicar a los no beneficiarios, tales como nombre y domicilio.

⁹ En 2002 existía el programa de Tecnificación del Riego como parte de la Alianza, que a partir de 2003 se asimiló al programa de Fomento Agrícola.

contrastantes: Oaxaca y Sonora ¹⁰ con objeto de tener la más amplia gama de aprendizajes metodológicos. Posteriormente, el programa de Desarrollo Rural se eliminó del diseño piloto debido a la gran cantidad de tipos de apoyo, y por lo tanto de tratamientos, de este programa. Por su parte, Oaxaca también fue eliminado del diseño inicial debido a su compleja institucionalidad y a las dificultades que presentaba para obtener información.

El diseño de evaluación cambió progresivamente al considerar la disponibilidad de información pre-pareo de los no beneficiarios y la homogeneidad del programa a evaluarse con respecto a los tipos de apoyo o tratamientos, de tal forma que el diseño experimental de evaluación terminó por aplicarse en la región de los Altos del estado de Jalisco al programa de mejoramiento genético. El principal aprendizaje logrado mediante este ejercicio piloto de evaluación fue que el alto costo marginal no se compensa por el beneficio marginal de contar con información más precisa sobre los impactos y su causalidad, sobre todo al tratarse de programas tan complejos como lo fue Alianza para el Campo. Además del costo elevado, la aplicación de este tipo de metodología no es factible para evaluaciones anuales. En el caso de la evaluación cuasi-experimental piloto, la encuesta se levantó a fines de 2004 y principios de 2005 a beneficiarios que recibieron los apoyos para mejoramiento genético en 2001, período en el que fue posible que se diera un proceso de maduración de la intervención.

b) 10 lecciones aprendidas de la evaluación sumativa de Alianza para el Campo

i) Sin duda, dada la complejidad del Programa de Alianza para el Campo, se realizaron mejoras significativas a lo largo de la experiencia de evaluación de este Programa. No obstante, no haber contado con información contrafactual, ya sea de línea de base o a través de grupos de comparación, fue una limitación importante para hacer una valoración rigurosa que se pudiera inferir sobre la causalidad de los impactos de las diferentes intervenciones de Alianza. Si bien antes de 2003 el diseño de Alianza cambiaba anualmente, a partir de 2003 se decidió mantener el diseño implícito en las Reglas de Operación, que habría representado una oportunidad para establecer una línea de base, como si se tratara de un programa nuevo, o bien una línea basal en cualquier momento de la operación de Alianza, después de 2003, a los beneficiarios que hubiesen recibido apoyos por primera vez y dar el carácter de contrafactual a esa información. El aprendizaje de la experiencia en la evaluación de Alianza indica la necesidad de contar con información contrafactual y que si bien lo ideal es levantarla desde el inicio del Programa, dicha información es útil incluso cuando se obtenga en plena operación del Programa y delimite claramente dos momentos.

ii) Si bien Alianza y sus programas sufrían frecuentemente cambios sustanciales en su diseño, en un programa de fomento como éste pueden visualizarse indicadores básicos de desempeño, que se mantienen vigentes en el largo plazo. El mejoramiento del ingreso y la competitividad de las cadenas productivas son ejemplos de dichos indicadores, con indicadores indirectos de desempeño tales como mecanización, capitalización, cambio tecnológico, etc. La lección derivada es, por lo tanto, que a pesar del cambio de los programas, es posible obtener información basal para indicadores estratégicos de medición de impactos. En ese sentido, el rigor que no tuvo la evaluación de Alianza en la medición de impactos y su causalidad, pudo haberse

¹⁰ Oaxaca y Sonora son estados con nivel relativo de desarrollo diferente.

tenido mediante el levantamiento de información contrafactual de indicadores estratégicos de pertinencia vigente en relación con su inherencia a cualquier programa de fomento rural.

iii) Prever la evaluación resulta importante para el diseño de los programas públicos, con la intención de incorporar indicadores de desempeño por lo menos al nivel del propósito y del fin ¹¹ que el Programa pretende contribuir. El número de dichos indicadores debe, sin embargo, ser acotado a fin de focalizar su seguimiento. Ello habría permitido no sólo incorporar desde el principio el levantamiento de información contrafactual, sino también focalizar los indicadores de desempeño del Programa en términos de su eficacia y eficiencia.

iv) Con respecto a la lección anterior, la definición de indicadores de desempeño permite darles seguimiento en evaluaciones sucesivas y generar bases de datos comparables entre los distintos ejercicios de evaluación.

v) En la definición de la temporalidad para la medición de impactos es importante considerar los tiempos de maduración de las inversiones dada su naturaleza. En el caso de la experiencia de evaluación de Alianza, la periodicidad anual para la medición de impactos resultó ser excesiva, ya que muchas de las inversiones otorgadas en apoyo a los beneficiarios aún no habían madurado al momento de la evaluación, que, por normatividad, debía realizarse anualmente. Por lo tanto, es recomendable definir, con base en criterios técnicos, la periodicidad de la evaluación de impactos a fin de que sea costo-efectiva y pertinente.

vi) La incorporación de la evaluación del Programa desde su diseño implica también la identificación de la información necesaria para medir los indicadores, tanto de gestión como de impactos. Esto a su vez significa que la toma de información relevante para evaluar el desempeño del Programa, es parte de la gestión del mismo.

vii) Las alternativas metodológicas para la medición más precisa de impactos y el establecimiento de su causalidad, como son los diseños cuasi-experimentales, encarecen significativamente los procesos de evaluación, especialmente si son anuales, y, en especial, al tratarse de programas complejos como Alianza. ¹² En este sentido, el costo que implica la precisión de la causalidad y medición del impacto generalmente es demasiado alto en relación con el beneficio derivado de ello.

viii) En el medio rural se presentan importantes limitantes para obtener información técnico-productiva confiable, debido a que no existen registros de costos de producción, especialmente de los pequeños productores. Por ello, es importante emplear metodologías e identificar fuentes más pertinentes para la obtención de información de campo, tales como estudios de caso o grupos focales, que, al integrarse con metodologías cuantitativas, permiten tener una mejor comprensión de los resultados y efectos del Programa evaluado.

¹¹ El fin y el propósito son términos empleados en la metodología de marco lógico para referirse a los objetivos en el diseño de un programa.

¹² En particular con relación a la cantidad de tratamientos y a la población cambiante de beneficiarios.

ix) Al no estar establecidos la evaluación y, en particular, los indicadores de desempeño desde el diseño del Programa, se dispersan los objetivos de la evaluación, que, a su vez, tienen como consecuencia la búsqueda de información que rebasa las necesidades de la misma. A menudo dicha información es además subutilizada. Para la estimación de impactos se debe recoger información congruente con los objetivos de evaluación y los indicadores predefinidos para tal efecto. Así, es necesario evitar la dispersión de esfuerzos y recursos, y limitar la recolección de datos a aquellos cuya utilización en la evaluación está claramente predefinida.

x) Los objetivos definidos formalmente para un Programa de diseño nacional y de operación local pueden diferir de los objetivos que los operadores persiguen en la práctica, quienes al final son los que deciden la asignación de los recursos y, por lo tanto, orientan el tipo de resultados que rinde el Programa. En ese sentido, el aprendizaje derivado es que los operadores participen en el diseño del Programa, ya que es a quienes finalmente va destinado.

b) Evaluación formativa (procesos)

La forma en que se implementan los programas, es decir, sus procesos o gestión, es un factor que determina en gran medida sus resultados e impactos finales. Más allá de la calidad y pertinencia del diseño, y de los objetivos planteados para cada Programa, las acciones de los operadores y de los responsables en la toma de decisiones de los distintos niveles son las que definen el impacto de la acción pública.¹³ Por lo anterior, a partir del ejercicio de evaluación de Alianza 2001, la evaluación de la gestión constituyó un ámbito central del análisis; año con año se avanzó en la metodología de evaluación de la gestión, y se distinguieron las siguientes etapas:

i) Evaluación de la operación del período 1998-2000. En una primera etapa, que va de la evaluación del ejercicio 1998 a la de 2000, el análisis realizado en las evaluaciones estaba compuesto principalmente por los siguientes ejes:

1) Evaluación de la operación. Incluía una descripción de los procesos de operación del Programa tales como planeación, difusión, elaboración y pago de solicitudes. Además de dicha descripción, se hacía un breve análisis del desempeño de cada proceso y se construyeron índices a partir de las encuestas a beneficiarios.

2) Valoración de la ejecución de los procesos operativos. Se hacía un breve análisis de los procesos con respecto a las funciones, tiempos y responsabilidades atribuidas por la normatividad a cada agente. Para ello se construían índices a partir de encuestas a funcionarios y otros actores.

3) Cuantificación de los avances en el cumplimiento de las metas físicas y financieras. Se establecían índices de eficiencia y eficacia ligados al ritmo de ejercicio de los recursos presupuestales con respecto a lo programado.

¹³ En el caso de Alianza, la gestión incluía la asignación de los recursos por parte de los operadores del Programa, que es una determinante importante en los resultados esperados.

4) Medición de la satisfacción de los beneficiarios de los programas con respecto a la calidad de la atención de los operadores, oportunidad y calidad de los apoyos, entre otros. Para ello se construían índices de satisfacción a partir de información obtenida mediante la encuesta a beneficiarios.

Las Entidades Evaluadoras Estatales (EEE) realizaban el levantamiento de encuestas para la evaluación de la operación en los estados para cada uno de los programas evaluados. Las fuentes básicas de información eran la opinión y percepción de actores que participaban en los programas (operadores, beneficiarios y líderes de organizaciones, entre otros) recabadas mediante las Encuestas a Beneficiarios y las Encuestas a Otros Actores (EOA). Con dicha información se conformaban bases de datos con las que se estimaban indicadores para las evaluaciones estatales y nacionales por programa.

La estructura de las encuestas permitía la construcción y agregación de indicadores, al generar datos duros para soportar la evaluación. Sin embargo, se observó que a partir de opiniones, las respuestas generalmente tenían un sesgo. En cuanto al beneficiario, el sesgo hacia la buena opinión con respecto a la operación se explica porque, independientemente de la oportunidad y calidad de la atención, los encuestados temían que una opinión negativa pudiera traducirse en dificultades para acceder nuevamente al Programa.

En lo que respecta al funcionario, opinar acerca de los procesos de gestión era de alguna manera ser juez y parte, y una mala calificación cuestionaría su propio trabajo. Así, las evaluaciones de la operación arrojaban en general indicadores de un buen nivel de desempeño global, aun cuando en la realidad se observaban importantes ineficiencias y dificultades operativas, como por ejemplo el retraso en el inicio de la operación y en los pagos a los beneficiarios.

Así que el carácter descriptivo y los sesgos de la información son las principales críticas a la evaluación de la operación realizada bajo la metodología antes mencionada. Además, dado que el enfoque y objetivo central era la verificación del cumplimiento de la normatividad, no se hacía un análisis crítico de la gestión que permitiera determinar relaciones de causalidad entre los problemas observados en determinados procesos y los resultados de los programas de Alianza. Asimismo, no se emitían recomendaciones concretas sobre acciones de mejora en los procesos a nivel estatal. Además, bajo este esquema, incluso en las evaluaciones nacionales no se contaba con elementos específicos para evaluar la gestión del Programa a cargo de las áreas centrales de SAGARPA, ya que el análisis se fundamentaba en información recabada en los estados y se orientaba al análisis de procesos estatales.

ii) Evaluación de procesos 2001-2003. La evaluación del ejercicio 2001 incluyó importantes modificaciones en su orientación de evaluación de la gestión. La evaluación estatal se estandarizó, al mejorar la construcción de indicadores para el análisis de los procesos de planeación, operación, arreglo institucional, difusión, gestión de solicitudes, otorgamiento de apoyos, evaluación y seguimiento, y se mantuvo el análisis de la satisfacción de los usuarios a partir de las encuestas a beneficiarios. A pesar del predominio de un enfoque descriptivo, se sentaron las bases para un análisis de mayor precisión de las áreas de oportunidad en la gestión.

A nivel de la evaluación nacional de los programas, se presentaron dos modificaciones sustanciales. Por una parte se incluyó, además de las fuentes estatales de información, un conjunto de visitas de campo a estados seleccionados, a fin de obtener información de primera mano y entrevistar directamente a los principales operadores de los programas. Con ello fue posible contrastar la información que arrojaron las encuestas y emitir recomendaciones con mejores fundamentos. Por otra parte, la evaluación brindó una visión de conjunto y permitió ir más allá de la verificación del cumplimiento de la normatividad al incorporar elementos de análisis de políticas públicas. Así, con mayores grados de libertad, se emitieron recomendaciones orientadas a mejorar la efectividad de la gestión del Programa con el objetivo final de incrementar los impactos de los recursos, y ya no simplemente el de medir los avances en metas o apego a indicadores establecidos en la norma.

En ese mismo año (evaluación 2001 realizada en 2002) se llevó a cabo la primera Evaluación Nacional de Procesos que, en lugar de limitarse al enfoque técnico productivo que prevalecía en las evaluaciones anteriores, introdujo una perspectiva de análisis organizacional y de implementación de política pública. Este enfoque permitió agregar grados de libertad al análisis, ya que sometió a consideración el propio diseño de Alianza, su orientación estratégica y la pertinencia del arreglo institucional, en lugar de limitarse a contrastar la realidad operativa con el deber-ser establecido en las Reglas de Operación.

Desde ese año y hasta la evaluación de 2003 se destacaron dos ejes de evaluación de la gestión: por una parte, la línea de evaluación nacional de procesos que se realizó por tres años consecutivos (del ejercicio 2001 al de 2003) y, por otra, la evaluación de procesos de cada programa. A pesar de ser dos niveles distintos de análisis, el enfoque de la evaluación nacional de procesos antes descrito constituyó la base sobre la que se diseñó la metodología para las valoraciones estatales y nacionales por programas, a partir de la evaluación de 2002. En opinión de evaluadores y usuarios, ese cambio de enfoque permitió que las evaluaciones ganaran en utilidad para la toma de decisiones, ya que permitían:

i) Realizar una valoración crítica sobre aciertos y oportunidades de mejora de los procesos más importantes de Alianza para el Campo que incidían de manera significativa en el impacto del Programa en el desarrollo agropecuario y rural.

ii) Aportar a los gobiernos federal y estatales recomendaciones para el mejoramiento del diseño y de los principales procesos de gestión de Alianza.

Asimismo, la inclusión de un enfoque de análisis continuo, al estudiar no sólo los procesos del año evaluado, sino al reportar observaciones sobre la operación de varios años, permitió emitir recomendaciones acerca de temas críticos de la gestión. Los elementos del enfoque que a partir de entonces orientaron la evaluación de la gestión en sus distintos niveles fueron:

i) Utilidad práctica. Para brindar a los funcionarios elementos de juicio que contribuyan a orientar la instrumentación de los programas de Alianza hacia una mayor eficacia en el logro de sus objetivos.

ii) Oportunidad de los resultados. Proporcionar información útil a los ámbitos de decisión, tanto federal como estatal, en varios momentos de la operación de los programas de Alianza.

iii) Análisis continuo. Brindar información sobre el desempeño de los programas de Alianza de forma continua y no sólo en el período del ejercicio al que corresponde la evaluación.

En la evaluación 2002, paralelamente al proceso de evaluaciones nacionales y estatales por programa, se realizó la segunda Evaluación Nacional de la Gestión que, si bien mantuvo una estructura y enfoque similar a la de 2001, se orientó a profundizar en el análisis de temas relevantes detectados en la evaluación anterior. Además, se focalizó el análisis transversal de la gestión de los tres grandes grupos de programas (Fomento Agrícola, Fomento Ganadero y Desarrollo Rural), al estudiar temas relativos al diseño, la planeación, el arreglo institucional, la operación, y el seguimiento y evaluación de los programas. Ello permitió detectar procesos críticos al interior de cada uno de esos temas, así como su relación con el desempeño de los programas.

Por su parte, la Evaluación Nacional de la Gestión 2003-2004 tomó como punto de partida el análisis de los problemas primordiales detectados en ejercicios anteriores. El principal resultado de dicha evaluación fue el análisis a profundidad y avance en el conocimiento de las causas y factores que explicaban los grandes problemas de implementación del Programa, que se resumieron en tres ejes:

1) Gestión pública descentralizada. Parte de las dificultades operativas del Programa se derivaban de la complejidad de su diseño, dado que participaban en su gestión un gran número de agentes (funcionarios estatales, federales y productores, entre otros) de distintas naturalezas, cuya coordinación constituía un reto.

2) Asignación de recursos. Los procesos estratégicos y operativos para determinar con precisión el destino de los recursos del Programa presentaban grandes áreas de oportunidad, a pesar de las lagunas, ineficiencias e inercias que impedían lograr una focalización sistemática y congruente con el impulso a líneas de política y objetivos de desarrollo del sector.

3) Aspectos críticos de la operación. Dada la complejidad del arreglo institucional, en el circuito operativo de Alianza se presentaron importantes ineficiencias tales como duplicación de acciones, lagunas y retrasos que se tradujeron en mayores costos operativos y menores impactos.

iii) Evaluación de la gestión 2004-2005. La Evaluación Nacional de la Gestión 2003-2004 fue el último ejercicio de esa naturaleza, ya que las modificaciones anuales a la gestión, vistas a nivel nacional y del Programa en su conjunto nacional, resultaron marginales, por lo que no se justifica un análisis anual. En ese marco, los temas de evaluación de la gestión constituyeron, desde el ejercicio 2004, los ejes centrales de la evaluación de la gestión de los programas, tanto a nivel nacional como estatal, y se consolidó la orientación de la evaluación de procesos hacia el análisis de la implementación de políticas públicas.

En preparación de la Evaluación de Alianza 2005 se realizaron ejercicios de valoración *ex post* de los distintos instrumentos empleados para la obtención de información relativa a la gestión. Con respecto a las Encuestas a Otros Actores (EOA), distintos usuarios, tanto a nivel nacional como estatal, reportaron que la información recabada mediante estas encuestas era poco utilizada en las evaluaciones. Uno de los principales problemas era que la mayor parte de la información se recibía de la opinión de diversos actores con respecto a los mismos temas, con lo que se obtenían distribuciones de frecuencias, pero no datos actuales de la situación de la gestión en cada Estado. Además, los usuarios consideraban que la relación utilidad/esfuerzo era baja; es decir, el tiempo invertido y el costo de búsqueda de información no se reflejaba en un nivel relevante de la utilización de la misma.

Así, para el ejercicio de evaluación 2005 se definió modificar el instrumento de obtención de información para la evaluación de la gestión de Alianza para el Campo. Por una parte, se diseñó una “Guía de Entrevistas a Otros Actores para la Evaluación Estatal” a fin de orientar a las Entidades Evaluadoras Estatales en las entrevistas abiertas a funcionarios que determinaran importantes. Con esto se eliminó la obligatoriedad del levantamiento de encuestas cerradas a un número predeterminado de actores que implicaba la EOA. El hecho de dejar abiertas las entrevistas permitió seleccionar mejor la temática y el tiempo destinado a cada actor para focalizar mejor los esfuerzos.

Por otra parte, se desarrolló la Cédula de Información Verificable sobre la Gestión Estatal de Alianza para el Campo, instrumento de levantamiento de información que permitía analizar la trayectoria y la situación actual de la gestión estatal a partir de la comprobación de hechos o de información factual sobre temas relevantes de la gestión de cada programa.

En opinión de los distintos usuarios, los cambios en los instrumentos fueron satisfactorios. Por una parte la Cédula captaba, más fácilmente, información de mayor utilidad que las antiguas EOA, ya que permitía obtener datos a partir de la fuente considerada como la más fiable para valorar la situación de la gestión en cada Estado. Por otra parte, afirmaban que la realización de una serie de entrevistas abiertas con otros actores era importante para corroborar y complementar la información mediante la Cédula recogida.

c) Ocho lecciones aprendidas de la evaluación formativa de Alianza para el Campo

i) En el diseño de los programas públicos es importante incorporar la evaluación como parte de las actividades de gestión de los mismos. En la medida de lo posible es necesario definir, previamente a la operación del Programa, la metodología e indicadores de evaluación de impactos.

ii) La evaluación debe ser parte de la gestión y, como tal, debe retroalimentar de manera clara y oportuna a los responsables en las tomas de decisión para definir la continuidad, modificación o incluso conclusión del programa evaluado.

iii) Continuamente la mayoría de los problemas no se encuentran en el diseño de los programas, sino en las acciones que los instrumentan. Así, incluso con un buen diseño, los programas pueden reportar bajos niveles de impacto debido a una deficiente instrumentación,

como fue el caso de Alianza para el Campo. Los impactos de un excelente Programa se esconden si se tienen dificultades operativas. Por ello, la evaluación de procesos cobra importancia pues puede aportar, de manera oportuna, elementos que retroalimenten la operación de los programas para introducir medidas correctivas que reorienten las acciones de los mismos.

iv) A pesar de su valor, la evaluación de la gestión tiende a subestimarse con respecto a la valoración de impactos. Ello se debe a que los impactos son las medidas con las que se promueven las acciones realizadas ante órganos fiscalizadores, e incluso ante la población en general, mientras que los resultados del análisis de la gestión tienen utilidad al interior de las áreas responsables como insumo para la instrumentación de acciones de mejora. De hecho, en el ejercicio de la evaluación estatal de Alianza, el mayor esfuerzo se dedicaba al levantamiento de la encuesta a beneficiarios para obtener la información para el cálculo de indicadores de resultados e impacto, en detrimento de la calidad de la evaluación de procesos, cuyos resultados habrían implicado la irrelevancia de la estimación de impactos ante los problemas detectados en la operación.

v) La evaluación de la gestión es más útil en la medida en que se extiende a un análisis de las políticas públicas, por lo que es necesario que la metodología contemple este aspecto y los evaluadores tengan formación en la materia.

vi) En el caso de la evaluación de programas descentralizados es importante considerar que la evaluación nacional no es la suma o agregación de las subnacionales. Existe en materia de gestión pública una distribución de atribuciones entre niveles de gobierno e instancias que debe evaluarse con enfoques, instrumentos y metodologías propias.

vii) En la evaluación de la gestión es importante, además de tener una valoración de los procesos en conjunto, focalizar el análisis hacia aquellos que inciden de manera significativa en los impactos del Programa. Así, para hacer recomendaciones útiles y realistas, es necesario conocer la estructura de incentivos e inercias de los distintos agentes, ya que éstos, y no los objetivos formales de los programas, son los que determinan los resultados.

viii) La observación directa de los evaluadores y las entrevistas abiertas con los operadores y beneficiarios constituyen las fuentes más importantes de información cualitativa para la evaluación de la gestión, aunque es importante complementarlas con el análisis documental y encuestas.

III. EVALUACIÓN DE POLÍTICAS PÚBLICAS PARA EL DESARROLLO RURAL SUSTENTABLE: EL CASO DE MÉXICO

El IICA ha desarrollado un instrumento útil para calcular los procesos de evaluación y dar seguimiento a los impactos territoriales de las políticas públicas, que estiman los indicadores de la tercera dimensión del Sistema de Monitoreo y Evaluación del Desarrollo Rural Municipal (SIMOM). Su objetivo es crear indicadores que permitan observar las diferentes dimensiones del desarrollo rural sustentable.

La intención es poder dar seguimiento a la combinación de variables de desarrollo, tanto las sociales como las económicas, ambientales y de política institucional. Cada una abarca un conjunto de componentes que pueden reflejarse en indicadores y datos para construir índices objetivos comparables espacial y temporalmente. Así, se tendría un indicador que contribuyera al proceso de toma de decisión sobre las políticas y al seguimiento de los comportamientos de sus impactos, referidos a espacios territoriales específicos.

Desde la perspectiva metodológica, la integración de indicadores en una sola función de medición debe observar ciertas consideraciones operativas y técnicas. En la referida a la selección de indicadores que deben formar parte de la función de síntesis, se deben cumplir requisitos de independencia, universalidad, oportunidad, temporalidad, pertinencia y relevancia. Posteriormente, se considera el tratamiento de las formas cuantitativas en que se expresan estos indicadores. Debido al sentido, la dirección, la cuantificación, su expresión numérica y sus marcos referenciales, la combinación de indicadores es una tarea especialmente delicada.

El modelo de estimación del Índice de Desarrollo Sustentable aporta una solución a ambos aspectos a través de un instrumento. Esta solución se alcanza con una elaboración de criterios para la selección de los indicadores que deben ingresar en la función de estado del sistema y un método estadístico de estandarización para establecer la comparabilidad y agregabilidad de los indicadores individuales. Igualmente añade un mecanismo para la introducción de consideraciones sobre las ponderaciones que cada analista puede asignar a alguna dimensión o componente en un caso determinado o en un territorio en particular.

Como complemento, se ha desarrollado un aplicativo que permite el manejo sencillo de las bases de información para crear y controlar el índice mediante un modelo denominado *Biograma*, que utiliza coordenadas radiales para mostrar una imagen de los equilibrios entre las diferentes dimensiones del desarrollo sustentable.

Este instrumento ha resultado particularmente útil para realizar análisis comparativos intertemporales y regionales, a partir de la posibilidad de establecer jerarquías ciertas y únicas de los territorios, con mediciones comparables y agregables.

En el caso de México, la definición de indicadores se realizó al tener presente la medición de los procesos de Desarrollo Rural Sustentable.

Los impactos se clasificaron en cinco grandes áreas de desarrollo: económico, ambiental, de cohesión social, cultural y de cohesión territorial. Así, se definieron indicadores para cada componente: en la dimensión de las políticas estratégicas, la matriz contiene 35 indicadores para visualizar el desarrollo de estas políticas sobre la inserción de los diferentes componentes que considera la Ley a nivel municipal, distrital y estatal. Dichos indicadores se distribuyen en los siguientes subprocesos: institucional (22), capacidades (6) y de financiamiento (7).

Para observar los procesos generados por estas políticas estratégicas, se elaboraron 39 indicadores, distribuidos en los siguientes subprocesos: socialización de la ley (3), planeación (9), gestión (9) y participación (18). Con los impactos o resultados esperados por estos procesos se generaron 43 subprocesos: económico (12), ambiental (8), cohesión social (15), cultural (4) y cohesión territorial (4).

En total, resultaron 117 indicadores a nivel de procesos y subprocesos para observar el desarrollo rural sustentable con enfoque territorial, en el marco de LDRS. Si se suman los 3 indicadores de las tres dimensiones y el indicador global sobre el desarrollo rural sustentable a nivel municipal, se tienen 121 indicadores generados por este modelo de gestión y evaluación para el desarrollo rural sustentable.

1. Sistemas públicos de información

Con la finalidad de aprovechar la información existente para el desarrollo del sistema, se realizó una revisión de las bases de datos del país. La lista siguiente de los sistemas de información que se consultaron corresponden a las instituciones consideradas por la Ley de Desarrollo Rural Sustentable Municipal, publicada en México en diciembre de 2001: Sistema de Información del Sector Rural (SISER), Sistema de Información Geográfica del Agua (SIGA), Monitor Agroclimático, Sistema de Información Ambiental y de Recursos Naturales (SNIARN), Sistema Unificado de Información Básica del Agua (SUIBA), Servicio de Información Agroalimentaria y Pesquera (SIAP), Instituto Nacional de Estadística, Geografía e Informática (INEGI), Servicio Meteorológico Nacional (SMN), Servicio Nacional de Información de Mercados (SNIM), dependiente de la Secretaría de Economía, Información en la Secretaría de la Reforma Agraria (SRA), Información en la SEDESOL, Instituto Federal de Acceso a la Información Pública (IFAI), Comisión Nacional Forestal (CONAFOR), entre otros.

2. Instrumentos de evaluación: Encuestas

Con base en la estructura de los indicadores, se definieron los cuestionarios a aplicar. Para cada indicador se elaboró una pregunta, y se conformaron los siguientes tres tipos de encuestas:

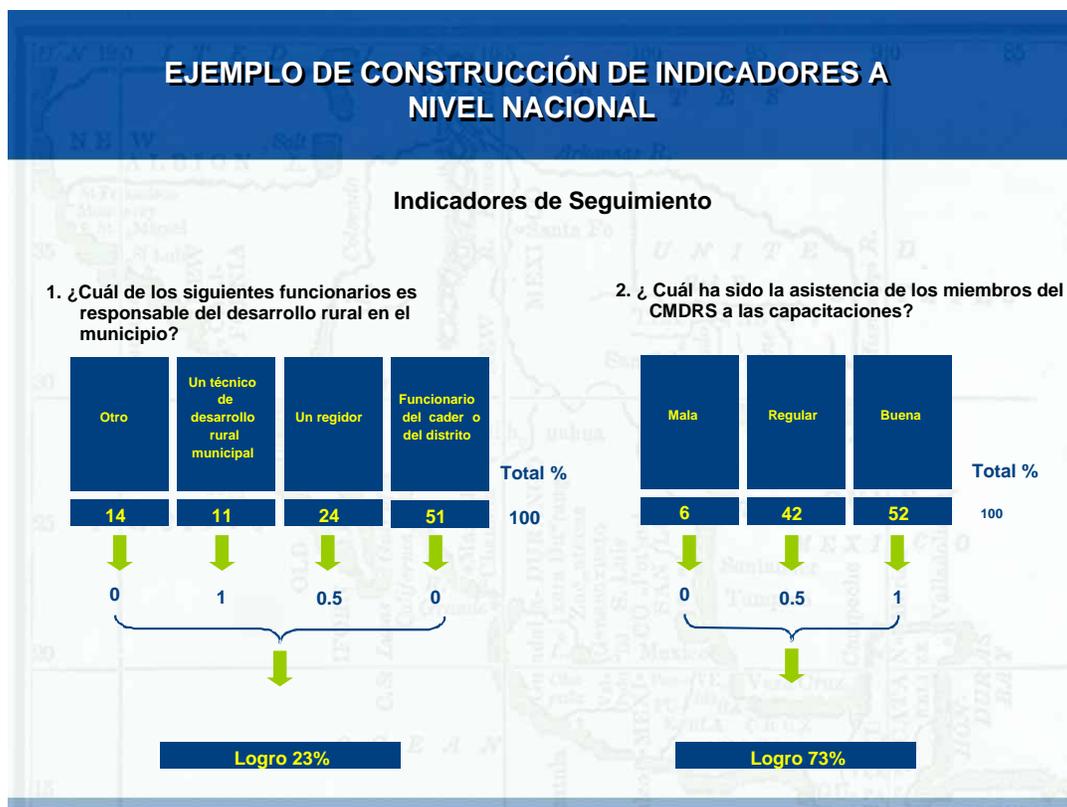
Coordinador del Consejo Municipal. Se aplica una encuesta para el *Coordinador del Consejo Municipal de Desarrollo Rural Sustentable (CMDRS)*.

Miembros del Consejo. Las encuestas para los *Miembros del Consejo Municipal de Desarrollo Rural Sustentable* se aplican a los líderes gremiales de organizaciones o representantes de organizaciones que integran el Consejo.

Funcionarios del ayuntamiento. La encuesta se aplica al funcionario del ayuntamiento involucrado en las tareas de planeación o en las del desarrollo rural, así como al presidente municipal o, en su caso, al regidor involucrado en los trabajos de la implementación de la Ley de Desarrollo Rural Sustentable en el municipio o en las tareas de planeación. De la misma forma, se administra al director agropecuario del ayuntamiento o, en su caso, al técnico adscrito a la presidencia municipal que estaba involucrado en las tareas del desarrollo rural o de planeación.

3. Estimaciones

Para la cuantificación de los indicadores se aplica una técnica de asignación de valores a las respuestas de las encuestas, como por ejemplo:



¿Cuál de los siguientes funcionarios es responsable del desarrollo rural en el municipio?

Las opciones de respuestas en la encuesta cerrada del Sistema son las siguientes: Técnico del DRSM, al que se le dio el valor de uno, el más grande, puesto que es el mejor resultado que se espera obtener en esta materia. Le sigue la opción de un regidor del ayuntamiento, al que se le asignó el valor de 0,5, ya que este funcionario todavía está considerado como una buena alternativa para impulsar el proceso de inserción de la ley en los municipios por actores que viven en él y que garanticen un mejor entendimiento de la problemática municipal. Como en el marco

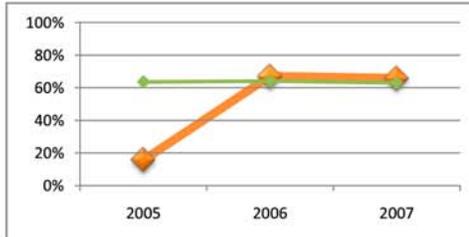
de la ley, se considera que éstos son los mejores funcionarios para dirigir el desarrollo rural en el municipio, a las demás opciones se les determinó el valor de cero. Entonces, el Sistema multiplica 11 por 1, y 24 por 0,5 para luego sumar los productos obtenidos y expresarlos en términos porcentuales, con un resultado de 23%.

A continuación se presenta un ejemplo de cómo se pueden apreciar los resultados: una comparación de indicadores del municipio de Charcas con el promedio de los indicadores de su estado, San Luis Potosí.

Siguiente →

Formalización de instituciones municipales para la promoción del desarrollo rural sustentable municipal

Constitución de Consejos Municipales de Desarrollo Rural

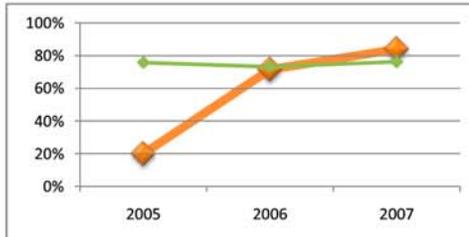


2005	2006	2007
16%	67%	66%

Variación	
05/06	06/07
Green	Yellow

Diferencia con SAN LUIS POTOSI		
2005	2006	2007
Red	Yellow	Yellow

Creación y fortalecimiento de unidades técnicas y de apoyo

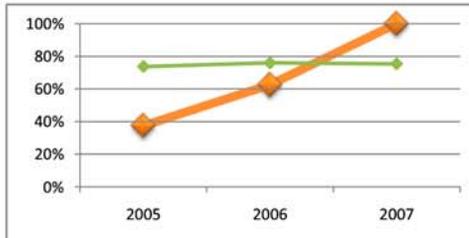


2005	2006	2007
20%	72%	84%

Variación	
05/06	06/07
Green	Green

Diferencia con SAN LUIS POTOSI		
2005	2006	2007
Red	Yellow	Green

Expedición de reglamentos operativos



2005	2006	2007
38%	63%	100%

Variación	
05/06	06/07
Green	Green

Diferencia con SAN LUIS POTOSI		
2005	2006	2007
Red	Red	Green

BIBLIOGRAFÍA

- Alkin, Marvin C. (2004), *Evaluation Roots: Tracing Theorists' Views and Influences*, Thousand Oaks, Londres, Nueva Delhi, SAGE Publications, febrero.
- Banco Mundial (2007), *How to Build M & E Systems to Support Better Government*, Keith Mackay, Independent Evaluation Group (IEG), Washington, D. C.
- _____ (2005), *Improving the World Bank's Development Effectiveness. What does Evaluation Show?* Operations Evaluation Department, Washington, D. C.
- _____ (2004), *Seguimiento y evaluación, instrumentos, métodos y enfoques*, Departamento de Evaluación de Operaciones, Washington, D. C.
- Campbell, D. y J. Stanley (1982), *Diseños experimentales y cuasi-experimentales en la investigación social*, Amorrortu, Buenos Aires, p. 16 (versión en inglés de 1966 publicada en Chicago).
- CEPAL (Comisión Económica para América Latina y el Caribe) (2005), “Evaluación de impacto”, *Serie Manuales*, N° 47, División de Desarrollo Económico, Santiago de Chile, noviembre.
- Chen, Huey-Tsyh (1990), *Theory-driven Evaluations*, SAGE Publications, California.
- Cook, T. (2005), *Métodos cuantitativos y evaluativos en investigación evaluativa*, Morata S. L., Madrid.
- _____ (1986), *Métodos cuantitativos y evaluativos en investigación evaluativa*, Morata S. L., Madrid.
- Chelimsky, E. y W. Shadish (1997), *Evaluation for the 21st Century. A Handbook*, SAGE Publications, California.
- Eisner, E. (1976), “Educational connoisseurship and criticism: Their form and functions in educational evaluation”, *Journal of Aesthetic Education* N° 10, julio-octubre.
- FAO (Organización de las Naciones Unidas para la Agricultura y la Alimentación) (2007a), “Report of the independent external evaluation of the Food and Agriculture Organization of the United Nations”.
- _____ (2007b), “Evaluations in FAO, new developments in evaluation” [<http://www.fao.org>].

- FAO/Proyecto FAO-SAGARPA (2007a), “Lineamientos metodológicos para la evaluación de la gestión de programas públicos derivados de la sistematización de la experiencia de evaluación de Alianza para el Campo”, México.
- _____ (2007b), “Sistematización de la experiencia en evaluación de impactos de Alianza para el Campo”, México.
- _____ (2006), “Sistematización de la experiencia del Proyecto Evaluación Alianza para el Campo”, México.
- Fetterman, D. (1994), “Steps of empowerment evaluation: From California to Cape Town”, *Evaluation and Program Planning*, N° 17.
- Grupo Independiente de Evaluación (2007), *Sourcebook for Evaluating Global and Regional*, Robert Picciotto, *Evaluation in the World Bank, Antecedents, Instruments and Concepts*, en E. Chelimsky y W. Shadish (1997), *Evaluation for the 21st Century*, OECD/DAC (2002), Evaluations and Results- Based Management.
- Guba, E. y L. Yvonna (1989), *Fourth Generation Evaluation*, California, SAGE Publications.
- House, E. (2000), *Evaluación, ética y poder*, 3ª Edición, Morata, S. L., Madrid.
- ILPES (Instituto Latinoamericano de Planificación Económica y Social) (2007), “Evaluación de programas”, Notas técnicas, Marcela Guzmán, Santiago de Chile, agosto.
- Ingram, Gregory K. (2003), “The challenges of development evaluation: Some conclusions in World Bank”, *Operations Evaluation Department, the first 30 Years*, OED, Washington, D. C.
- King, J. A. (1995), “Involving practitioners in evaluations studies: How viable is collaborative evaluation in schools?”, en Cousins, J. B. y L. M. Earl, *Participatory Evaluation in Education: Studies in Evaluation Use and Organizational Learning*, Falmer, Londres.
- MacDonald Barry, G. K. Verma y C. Bugley (ed.) (1979), *Racism and Educational Evaluation, Evaluation and Identity*, Londres.
- Patton, Michael Q. (1997), *Utilization-Focused-Evaluation: The New Century text*, 3rd Thousand, SAGE, Oaks, C. A.
- Provus, Malcolm (1971), *Discrepancy Evaluation*, Berkeley, California.
- Riecken, H. y B. Robert (1975), *Social Experimentation: A Method for Planning and Evaluating Social Intervention*, Academic Press Inc., Londres.
- Scriven, M. (1967), “The methodology of evaluation”, *Perspectives of Curriculum Evaluation*, Rand McNally, Chicago.

- SIDA (Swedish International Development Cooperation Agency) (2007), *Looking Back, Moving Forward, Evaluation Manual*, Stefan Molund, Göran Scholl, 2ª edición revisada.
- Stake, R. (2007), *Investigación con estudio de casos*, 4ª edición, Morata S. L., Madrid.
- Stufflebeam, Daniel L. y Anthony J. Shinkfield (2007), *Evaluation Theory, Models, and Applications, Chapter 1*, Western Michigan University, abril.
- Suchman, Edward (1967), *Evaluative Research: Principles and Practice in Public Service and Social Action Programs*, Russell Sage, Nueva York.
- Tyler, R. (1942), *Eight-year Study of Secondary Education*, Progressive Education Association, Universidad Estatal de Ohio.
- Weaver, L. y C. Bradley (2004), “Unpacking the participatory process”, *Journal of Multidisciplinary Evaluation*, N° 1, Michael Scriven y E. Jane Davidson (eds.), Kalamazoo, Michigan, The Evaluation Center, Western Michigan University, octubre.
- Weisner, E. (s/f), *Evaluations, Markets, and Institutions in the Reform Agenda of Developing Countries*, en E. Chelimsky y W. Shadish.
- Weiss, C. (1985), *Investigación evaluativa: Métodos para determinar la eficiencia de los programas de acción*, 5ª reimpresión, Editorial Trillas, México.
- Wholey, J. (2004), *Handbook of Practical Program Evaluation*, 2ª edición, Jossey Bass, California.

Anexo 1**ENLACES SOBRE EVALUACIÓN**

<http://www.oecd.org/puma>

OECD/Public Management Service, Improving Evaluation Practices: Best Practice Guidelines for Evaluation and Background Paper, 1999, the guidelines identify key.

<http://www.mande.co.uk/docs/MandENEWS2002progressreport.doc>

[http://www.uneval.org/index.cfm?module=Bookmarks&page=BookmarkList&BookInfrastructure, Agriculture, and the Environment marksHome=1](http://www.uneval.org/index.cfm?module=Bookmarks&page=BookmarkList&BookInfrastructure,Agriculture,and,the,Environment,marksHome=1)

W.K. Kellogg Foundation, Evaluation Handbook, 1998 (<http://www.WKKF.org/>).

Japón Evaluación (<http://www.jica.go.jp/english/evaluation/guides/guideline.html>).

Página del UNDP (<http://www.undp.org/eo/>).

Peter Rossi, Freedman & Wright. Evaluation a systematic approach, 1979.

Shadish, W., Cook, T., & Leviton, L. (1991). *Foundations of program evaluation: Theories of practice*, Newbury Park, CA: SAGE.

United Nations Evaluations Group ([http://www.uneval.org/index.cfm? module=UNEG&Page=UNEGDocuments](http://www.uneval.org/index.cfm?module=UNEG&Page=UNEGDocuments)).

USA (<http://dec.usaid.gov/partners/evalweb/resources/evalfour.cfm>).

www.oecd.org (website of the OECD/DAC Network on Development Evaluation).

www.worldbank.org (website del Banco Mundial con su Independent Evaluation Group, IEG. Contains a wealth of useful publications).

www.eldis.org

www.alnap.org (the Active Learning Network for Accountability and Performance in Humanitarian Action—an International Interagency).

www.policy-evaluation.org (wwwVirtual Library: Evaluation).