

PROBLEMAS ESPECIALES DE LA MEDICION DEL IMPACTO DE PROGRAMAS EN LOS PAISES EN DESARROLLO

John W. Townsend,¹ W. Timothy Farrell¹ y Robert E. Klein¹

Introducción

La aplicación de técnicas de evaluación en los países en desarrollo se ha extendido rápidamente durante los últimos años, si bien el número absoluto de programas de desarrollo social y económico que explícitamente contienen componentes de evaluación sigue siendo aún relativamente pequeño. Los ejemplos varían desde la evaluación de pequeños proyectos educativos en zonas rurales de Guatemala hasta programas regionales de gran envergadura en el Brasil. Los gobiernos no solo se interesan por evaluar los beneficios de estrategias de desarrollo adoptadas desde antiguo (por ejemplo, reforma agraria o asistencia técnica) sino también por mejorar los métodos de ejecución de proyectos ya existentes o experimentales (v.g., atención rural de salud, participación de la comunidad en la planificación del desarrollo, etc.).

Indudablemente, esta reciente y rápida expansión de las investigaciones evaluativas se debe en parte a que las entidades financiadoras insisten cada vez más en que los programas que ellas apoyan contengan componentes de evaluación. Además, la generación actual de políticos y administradores de programas suele estar mejor preparada que sus antecesores, y es también más consciente de los usos y beneficios potenciales de la evaluación de programas. Ellos ven la evaluación como parte integrante del proceso que aporta información para mejorar la adopción de decisiones y la planificación. Al utilizar la evaluación en los programas sociales, no sólo están determinando lo que sobre el papel son políticas razonables, sino también descubriendo los medios de convertir estas políticas en operaciones viables, a la vez que ensayan sus efectos en los programas.

En nuestra exposición sobre la investigación evaluativa, empleamos tres términos ampliamente usados en este campo (1). La *evaluación de procesos* se refiere al grado en que se ejecutó el programa de conformidad con su diseño inicial y según las pautas establecidas. Por ejemplo, ¿recibió el grupo o grupos tratados la intervención como se planeó inicialmente y en el grado especificado? La *evaluación de impactos* sigue a la evaluación de procesos y se refiere al grado en que el programa o intervención modificó la situación para la cual fue diseñado. La *evaluación global* se refiere a los estudios que contienen componentes tanto de la evaluación de procesos como de impactos.

En las secciones que siguen trataremos principalmente de algunos problemas operacionales, técnicos y éticos que son importantes para la evaluación en los países en desarrollo. Por último, comentaremos algunas de las oportunidades y necesidades especiales de la investigación evaluativa en esos países.

Problemas de estrategia asociados con la ejecución de programas en países en desarrollo

Son dos los problemas especiales de estrategia que plantea la ejecución de programas en los países en desarrollo: elección del momento oportuno para la intervención y divergencia entre los planes y las operaciones. Estos problemas deben ser abordados en todos los proyectos, cualquiera que sea su objetivo. Lo que hace esto particularmente agudo es que los científicos sociales comúnmente han limitado su participación en la evaluación de programas sociales a cuestiones de diseño experimental y medición de resultados (entre las recientes excepciones en Estados Unidos están el experimento sobre seguros de enfermedad (2) y el experimento sobre incentivos al trabajo para graduados (3)).

Recientemente algunos evaluadores (4) han aboga-

¹ Instituto de Nutrición de Centro América y Panamá.

do por un estudio más cuidadoso del proceso de ejecución de programas. Con frecuencia la ejecución de proyectos de intervención debe planearse de acuerdo con el ciclo anual de los participantes en el programa. Esto es especialmente cierto en cuanto a gastos, ingresos, y condiciones climáticas entre la población rural de los países en desarrollo, ya que estos factores pueden fluctuar grandemente durante el año. La disponibilidad de fondos y alimentos es mayor en las familias durante los períodos de cosecha y poco después. Si algún material nuevo o algunas prácticas llaman la atención de los pobladores de las aldeas durante este período, es más probable que los destinatarios de los programas estén más dispuestos a invertir dinero y tiempo en el producto o servicio que cuando los recursos de la cosecha anterior están ya casi agotados y las decisiones futuras de gastos deben esperar a la nueva cosecha. Además, los pequeños agricultores están mucho más ocupados durante ciertos períodos del año que en otros, y pueden emprender proyectos con mucho más éxito durante aquellos períodos en los que cuentan con más tiempo disponible. Los proyectos que se inician demasiado tarde en el ciclo agrícola puede que nunca sean completados, y podrían predisponer a los beneficiarios contra la participación en proyectos futuros (5). Los ciclos climáticos también pueden afectar el éxito de los proyectos. Por ejemplo, la asistencia a los programas de capacitación en el trabajo puede disminuir mucho durante la estación lluviosa debido a las dificultades para viajar y al recrudecimiento de las enfermedades gastrointestinales y respiratorias entre los pobres de las áreas rurales durante este período.

Otro aspecto de la elección del momento oportuno se relaciona con la ejecución de programas en gran escala. A menudo, cuando un programa grande (un programa de extensión agrícola, por ejemplo) se introduce a nivel nacional o regional, no se dispone de un grupo apropiado de comparación. Los análisis en serie cronológica podrían ser útiles para la interpretación de resultados, pero únicamente si se satisfacen las condiciones de ejecución. Primero, los programas nuevos deben introducirse en forma rápida, ya que los programas de aplicación gradual pueden producir cambios que muchas veces no se pueden distinguir a nivel nacional de las tendencias seculares a largo plazo. Es solamente el cambio relativamente brusco y coincidente en la producción agrícola, la dieta doméstica, y el estado nutricional y de salud de las familias de los pequeños agricultores el que puede infundir cierta confianza en la inferencia causal de que las actividades de extensión conducen a las mejoras mencionadas.

Además, para evitar efectos sutiles de regresión, la ejecución de un programa a nivel nacional o regional no debe hacerse a continuación de una baja grande en los indicadores de desarrollo. Una medida correctiva introducida de inmediato en respuesta a un problema agudo, aunque administrativamente atractiva, podría

ser difícil de distinguir del retorno usual a la tendencia normal manifestada por indicadores inestables (6).

El segundo problema general en la ejecución de programas en los países en desarrollo es el desajuste entre las intervenciones en la forma en que fueron planeadas inicialmente y su ejecución real. Cuando la ejecución de un programa está obstaculizada por la escasez de recursos, problemas de comunicación y de transporte, o algún otro problema de orden logístico, se debe dedicar atención muy especial al seguimiento de la prestación real del tratamiento (es decir, a la evaluación del proceso). La literatura sobre desarrollo rural está llena de ejemplos de intervenciones "fantasmas" (por ejemplo, clases no impartidas, publicaciones no distribuidas, equipo perdido o robado, etc.). En una reciente evaluación de un programa radiofónico de nutrición en Nicaragua, los investigadores descubrieron que los dos minutos contratados con las estaciones de radio no se estaban emitiendo. Además, esta falta de cumplimiento estaba relacionada directamente con la potencia, la complejidad tecnológica y la localización de la emisora. Las emisoras rurales pequeñas simplemente no respondían a los planes de intervención (7).

Es obvio que para todo tipo de evaluación es imperativo que el investigador pueda determinar si las intervenciones fueron aplicadas en la forma planeada. Aun cuando las intervenciones en los países en desarrollo con frecuencia son relativamente débiles en comparación con la magnitud de los problemas, las evaluaciones centradas en la fuerza y el momento oportuno de la intervención pueden servir para poner a prueba las alternativas de política. Siempre que la distribución, la magnitud y el momento de la intervención son factores cruciales, la evaluación debe incluir una serie de mediciones del proceso. Si la variación en la ejecución es independiente de las características de los participantes, pueden obtenerse valiosos datos sobre la sensibilidad de la intervención a distintas formas de ejecución. Por ejemplo, en los proyectos que requieren la participación de la comunidad, es útil saber qué cauces locales de influencia son más eficaces para diseminar la información y tienen mejor acogida que otros. En un proyecto de salud en la India, los empleados tropezaron con dificultades en un plan de evacuación hasta que los dirigentes religiosos de la comunidad dieron su aprobación. Este tipo de información puede sugerir mejoras operacionales sencillas, más que cambios básicos de la política.

Otro problema estratégico básico, muchas veces asociado con la labor de evaluación, son las tasas diferenciales de deserción de los participantes en el programa. Estas diferencias a menudo se presentan entre los grupos de intervención y los de control, a causa de diferencias de motivación o capacidad para continuar en el estudio (8), pero también pueden ser el resultado directo de la intervención misma.

Por lo general, el primer tipo de problemas de deserción se relaciona con abandonos en el grupo testigo.

Por ejemplo, es fácil imaginar tasas más altas de emigración debidas al rápido crecimiento de la población y, consecuentemente, a la falta de disponibilidad de tierras y de oportunidades de empleo en las comunidades que sirvan de grupo testigo en los programas de desarrollo rural. Las familias testigo pueden emigrar, ya sea porque su comunidad simplemente no cuenta con los beneficios del desarrollo o porque el acceso a beneficios no asociados con el programa de desarrollo es limitado. Aun cuando incluir más sujetos testigos puede ser relativamente fácil, puesto que su participación suele ser menos costosa, ello no resolverá los problemas de deserción porque lo que facilita la evaluación es la equivalencia de los grupos a lo largo del experimento. Si la deserción diferencial se debe a diferencias relacionadas con la propensión individual a emigrar de los sitios del estudio, el problema de la deserción podrá también verse como un tipo de autoselección que amenaza aún más la generalizabilidad de los resultados.

Cuando la deserción es producto de la intervención misma, la situación se torna más compleja. Un resultado de la mayoría de los proyectos de desarrollo es la introducción de ideas novedosas, aptitudes y estrategias para resolver problemas. Este desarrollo de las capacidades individuales unido a la atracción por la vida urbana (ganancias económicas, mejores escuelas y servicios médicos y nivel de vida más alto, etc.) puede contribuir a la "fuga de cerebros" de las comunidades rurales pequeñas. En otras palabras, los más propensos a aceptar las innovaciones y más activos en el proceso de cambio, son los que emigrarán con más probabilidad. En este caso la deserción toma las características de una variable dependiente. Una de las soluciones podría ser el seguimiento de los emigrantes, pero esto también plantea problemas. El principal obstáculo radica en la dificultad de localizar a la gente, especialmente a los miembros de poblaciones pobres que pueden tener una elevada movilidad. Los sesgos debidos a la deserción pueden ser muy significativos, puesto que los emigrantes más fáciles de localizar probablemente sean los que muestran haber tenido más éxito, tanto por su estabilidad social como porque los que han fracasado en el medio urbano estarán menos dispuestos a revelar su situación presente.

Sin lugar a dudas, no existen soluciones fáciles al problema de la deserción aunque se han propuesto varias (6, 8-10), como el pago por cooperar en las intervenciones, para que haya ganancias netas tanto para el grupo de intervención como para el testigo, y la mejora de la oferta de trabajo en los poblados pequeños para reducir la emigración. Esencialmente, esto equivale a comparar dos tratamientos activos en vez de comparar un tratamiento con un testigo. El punto más importante con respecto a la deserción es que la comparabilidad entre los grupos de intervención y los testigos es en último término mucho más importante que mantener el tamaño absoluto de los grupos.

Otro obstáculo técnico al desarrollo de la capacidad

de evaluación es la calidad de las instalaciones disponibles para el tratamiento de datos. Los problemas de análisis multivariante y las presiones de los planificadores de programas para obtener una retroalimentación rápida obligan a los evaluadores a recurrir a las técnicas más recientes de cálculo electrónico. Sin embargo, la gran escasez de personal de programación capacitado, computadoras y programas de computación, con frecuencia, obliga a los evaluadores a usar instalaciones de tratamiento de datos fuera del país. Aunque a menudo es necesario, esto puede acarrear consecuencias desafortunadas. La exportación de los datos para su análisis agudiza la dependencia del país. También aísla al personal de programas y administrativo de las actividades de análisis de datos e impide la formación de nacionales en las técnicas correspondientes.

Consideraciones operacionales en la planificación y ejecución de la evaluación

En los proyectos de evaluación en los países en desarrollo, algunas cuestiones operacionales exigen especial atención, éstas son los problemas de selección de la muestra, lenguaje, adecuación cultural, relaciones entre el personal del programa y el de evaluación, y financiamiento de los programas de investigación evaluativa. Por supuesto, existen otros, como la brecha entre teoría social, planeamiento de programas y ejecución, que también son importantes, pero, en el área creciente de evaluación de programas sociales en los países en desarrollo, los citados anteriormente parecen ser los que más necesitan de un examen minucioso.

Selección de la muestra

Quizás el criterio más importante para la selección de la muestra es la generalizabilidad de los resultados experimentales (11). Si los resultados obtenidos no pueden extenderse a otras personas y situaciones donde fuese aplicable un programa social similar, será poco lo que de importancia inmediata se aprenda de la intervención. Debido a que la naturaleza de la investigación evaluativa implica consecuencias políticas inmediatas, nunca se dará demasiada importancia a la representatividad de la muestra y a la generalizabilidad de los resultados. Los individuos o grupos seleccionados para participar en la evaluación de programas no deben tener características que los hagan inapropiados para el tratamiento o los impidan utilizarlo. Así, por ejemplo, los pequeños agricultores cuyas prácticas agrícolas están restringidas por el tipo de tenencia de tierra (arrendamiento o propiedad) no se deben elegir para que participen en la evaluación de un programa a largo plazo de mejoramiento de pastos. Debido a que los agricultores que no son dueños de la tierra tienden a cambiar de fincas cada

pocos años, sus reacciones a la intervención diferirían probablemente de las de la población que se desea estudiar (es decir, pequeños agricultores con intereses a largo plazo en determinadas parcelas de tierra).

Lo dicho no significa que debería evitarse la participación de casos extremos en proyectos de desarrollo "no investigativos". Por el contrario, un beneficio secundario del proceso de selección de muestras en los países en desarrollo es la identificación de individuos especialmente necesitados de ayuda por parte de los organismos de servicios.

Una segunda consideración general es que los participantes en la muestra no deben seleccionarse únicamente sobre la base de su deseo de cooperar, su disponibilidad, o su experiencia en programas de desarrollo. La menor generalizabilidad de los resultados obtenidos con muestras de voluntarios se suele denominar selección por interacción del tratamiento (12) y ocurre cuando la muestra de voluntarios es más susceptible al cambio social, está más inclinada a cooperar en programas comunales, o está más capacitada para aprovechar la oportunidad que el promedio de la población. Sin embargo, este problema no podrá evitarse cuando la participación en programas experimentales entrañe riesgos relativamente altos o cuando la participación en la intervención misma sea contraria a las normas culturales tradicionales. Por ejemplo, puede ser necesario recurrir a voluntarios en un programa experimental de comercialización agrícola entre agricultores marginales. La resistencia de la mayoría de éstos a abandonar las prácticas comerciales tradicionales puede no ser una indicación de que no están interesados en los potenciales beneficios económicos del programa, sino sólo que consideran que las consecuencias del fracaso serían demasiado graves. En zonas donde las cosechas son pequeñas y están a merced de plagas y accidentes meteorológicos, las estrategias de intervención, que buscan llevar hasta el máximo la generalizabilidad, deben minimizar los riesgos adicionales asociados con la participación.

La disminución del riesgo podría adoptar varias formas. Estas van desde la aplicación gradual del programa, de manera que los riesgos asociados con cada cosecha sean relativamente pequeños, hasta la asignación de presupuestos suficientes para proteger los ingresos de los colaboradores si acaso la cosecha fuese muy mala.

Otra consideración es que algunos subgrupos quizás valoran positivamente las innovaciones y el cambio. Por ejemplo, las comunidades que han tenido contacto prolongado con agentes de extensión y han experimentado los beneficios de la innovación técnica, estarán más dispuestas a continuar los programas de experimentación. El hecho de que una técnica sea nueva puede ser razón suficiente para experimentarla, y tal vez para adoptarla. En cambio, para individuos de zonas rurales de países en desarrollo, las innovaciones y el cambio en sí tienen menos atractivo. Si un individuo está condicionado para ver lo nuevo con escepti-

cismo y cautela a causa de las normas culturales, es probable que no coopere (13, 14). En tales casos, el menor uso de muestras de voluntarios probablemente reduciría la generalizabilidad de los resultados del programa.

Otro aspecto un tanto diferente del problema de los voluntarios son los proyectos en que la toma de decisiones comunitaria es parte integrante del modelo del desarrollo. En este caso, la selección de la intervención por la comunidad es la esencia del proyecto. La literatura sobre evaluación ofrece poca orientación a este respecto, puesto que los estudios de intervención generalmente han involucrado la administración diferencial de servicios públicos, más que el cambio estructural que implica la autodeterminación.

En definitiva, el problema de selección de muestras se reduce a equilibrar la necesidad de viabilidad logística y el deseo local de cooperar (10). En todo caso, los evaluadores en los países en desarrollo deben tener cuidado de identificar las características socioculturales de sus muestras y, cuando consideran necesario sacrificar la representatividad a otras conveniencias, deben informar a los posibles usuarios acerca de las limitaciones de la muestra.

Problema del lenguaje

Cuando el personal de evaluación no proviene del grupo en estudio, surgen problemas comunes a todas las investigaciones transculturales. Entre los más agudos está la ignorancia o insensibilidad del evaluador a las diferencias de lenguaje o dialecto, así como la pertinencia e idoneidad culturales de las mediciones de procesos y resultados usadas para evaluar la eficacia de los programas. En el examen que seguidamente se hace de estos problemas, presentaremos ilustraciones de cómo pueden impedir los esfuerzos de investigación, sugiriendo algunas estrategias para resolverlos.

La comunicación es una cuestión compleja, puesto que está entrelazada con otros importantes problemas de investigación transcultural, y abarca tanto la comunicación verbal como la no verbal. En esta sección se aborda principalmente la comunicación verbal, aun cuando se mencionan algunos de los factores no verbales más críticos (15, 16).

El problema básico es la mezcla de patrones de comunicación del personal de evaluación, del personal del propio programa, y de la población. La fluidez en el idioma nacional de un país muchas veces no es suficiente. Por ejemplo, un hispanohablante tal vez no domine las lenguas indígenas e, incluso en español, las diferencias en el habla regionales o de clases sociales pueden causar problemas de comunicación. El evaluador debe estar preparado para enfrentarse con los diferentes problemas de lenguaje en su trato con diversos grupos.

El evaluador, el gobierno y la comunidad científica. Aun cuando muchos científicos y funcionarios guber-

namentales de países en desarrollo hablan con fluidez varios idiomas, el evaluador debe hablar con suficiente fluidez el idioma nacional del país donde trabajará. En caso contrario, no solo se dificulta la comunicación con el personal del país huésped, sino que también, según el idioma en que rinda sus informes, será menor la probabilidad de que los resultados de la evaluación sean utilizados.

El evaluador y el personal del programa y de evaluación. Cuando se hace necesario el uso de entrevistas extensas en las encuestas para evaluar un programa, las deben llevar a cabo entrevistadores que hablen el "lenguaje de campo", o sea, el idioma cotidiano de la población. Los supervisores de los entrevistadores quizás tengan que ser bilingües para servir de puente entre éstos y el personal superior de evaluación, quien puede que ignore el "lenguaje de campo". Como es esencial una comunicación clara de las tareas, procedimiento de acopio de datos, etc., puede ser necesario al personal de evaluación adquirir, para su trabajo, cierto conocimiento del idioma en que se desarrollarán las operaciones del programa y las entrevistas de evaluación, especialmente si no se logra conseguir supervisores bilingües.

El evaluador y el idioma de la comunidad. El idioma de los destinatarios del programa constituye un factor complejo e importante. Es posible que surjan problemas de comunicación al nivel de la comunidad, no sólo debido a los distintos grupos lingüísticos, sino también a causa de los distintos dialectos, expresiones coloquiales y pronunciación local. Aun cuando los participantes del programa digan que saben hablar el idioma nacional, tal vez sea a un nivel insuficiente para una verdadera entrevista. Por ejemplo, en un estudio sobre la disposición a postergar el disfrute de bienes materiales, las preguntas exigirían el uso del subjuntivo presente y futuro. Mediante pruebas preliminares, se pudo establecer que esta forma verbal no era bien comprendida, hasta el punto de que varios informantes se sintieron engañados, ya que el investigador no había cumplido su "promesa" de entregarles tres vacas al cabo de un año. En estas circunstancias, quedó claro que, aun cuando los informantes hablaban el idioma nacional, lo hacían a un nivel inadecuado para realizar entrevistas complejas. En este caso se hizo necesario que el investigador aprendiera el lenguaje de campo apropiado y las frases indispensables para obtener datos sobre conceptos de esta naturaleza.

Aunque sería lo ideal, cabe preguntarnos hasta qué punto es realista que los evaluadores conozcan bien el idioma nacional y tengan por lo menos un conocimiento del lenguaje de campo suficiente para la realización de su trabajo. En una sociedad con múltiples culturas y grupos lingüísticos, la población estudiada puede incluir varios de estos grupos. ¿A qué compromiso se podría llegar que lleve al máximo la calidad del resultado de la investigación y reduzca al mínimo el tiempo y los gastos necesarios para dominar una serie de len-

guas y dialectos? Como mínimo, el investigador debe hablar fluidamente el lenguaje nacional y se debe contratar, cuando proceda, personal de evaluación completamente bilingüe. Si el evaluador no está familiarizado con el lenguaje de campo, es posible que sea necesario contratar los servicios de un traductor para facilitar la comunicación.

Como el lenguaje hablado y escrito no constituyen el único medio de comunicación, haremos una breve mención de la comunicación no verbal. Aunque no es razonable esperar que los investigadores se tornen multiculturales, conviene saber el significado de ciertos tipos de comportamiento no verbal. Algunos de los más importantes en Latinoamérica son, por ejemplo, la proximidad, el contacto visual y el contacto táctil. Entre las poblaciones indígenas de las tierras altas de Guatemala, una buena manera de poner fin a una entrevista, antes de que haya empezado, es una mirada directa y penetrante a los ojos del entrevistado y un "masculino" apretón de manos. Esa conducta connota una marcada falta de respeto.

Adecuación cultural de las medidas de evaluación

No menos importante que el problema del lenguaje es la adecuación cultural del proceso evaluativo y de las medidas de resultados. Por ejemplo, si se desea considerar la propiedad de tierra cultivable como una variable de resultados en un programa de desarrollo, le corresponde al investigador determinar, mediante estudios etnográficos, los diversos aspectos de la tenencia de la tierra en la población estudiada. Si las prácticas de tenencia de la tierra tradicionales son incompatibles con la propiedad de tierras cultivables (por ejemplo, tenencia comunal no formal) puede que el concepto no sea un indicador apropiado del desarrollo. Por ejemplo, en muchos países en desarrollo, los habitantes rurales toman a mal que sus vecinos adquieran alguna propiedad, porque sus normas tradicionales de compartir y de reciprocidad los llevan a pensar que nadie puede prosperar sin apropiarse de algo de su vecino (17). De manera similar, si la incidencia de infecciones urinarias en las mujeres embarazadas es una de las preocupaciones de un programa de salud rural, pero las normas culturales de pudor impiden formar una muestra aleatoria de mujeres embarazadas para efectuar las pruebas apropiadas, puede ser necesario abandonar ese indicador de salud o modificar la muestra para centrarse en pacientes hospitalizadas con otros padecimientos ginecológicos.

Las mediciones deben estar acordes con la población estudiada. Si se desea examinar la aceptabilidad de un nuevo cereal híbrido, los aspectos por los cuales se valore el grano deben ser culturalmente pertinentes, especialmente en cuanto a su descripción y evaluación (v.g., el maíz no sólo debe ser agradable al paladar, sino que también debe tener propiedades adecuadas para hacer tortillas).

Para evaluar la aceptabilidad de la mayoría de los

programas de intervención deben emplearse instrumentos culturalmente apropiados que midan la aceptabilidad según los valores semánticos indígenas. Por ejemplo, se pueden averiguar las preferencias utilizando instrumentos como las técnicas semánticas diferenciales, las pruebas de percepción temática modificadas o las técnicas de escala multidimensional (18-20). Tales medidas tienen la ventaja de ser comprendidas con relativa facilidad por los informantes y de tener las características de juegos, reduciendo así el aburrimiento y la fatiga del informante (21). Sin embargo, una desventaja, que acompaña a tales medidas en los países en desarrollo, es que requieren notable refinamiento por parte del entrevistador, una estandarización rígida, y considerables aptitudes de análisis estadístico.

El desarrollo de medidas culturalmente apropiadas con frecuencia requiere alguna investigación etnográfica antes de diseñar las medidas evaluativas, a fin de elaborar un marco operacional que encaje en la realidad cultural. A menudo, dicho marco tendrá que ser lo bastante sensible para que refleje las diferencias entre diversas comunidades. En algunos casos esto requerirá modificar las categorías operacionales, sin que por ello se cambie el dominio conceptual. Por ejemplo, los materiales de construcción de viviendas pueden utilizarse como índice de riqueza relativa dentro de una comunidad (22-24). Aunque esto parezca ser un buen signo de riqueza en algunas comunidades, también puede sufrir de falta de generalizabilidad, o sea que un material barato en una comunidad, por ejemplo, la madera, puede ser más caro en otra. Por lo tanto, es fundamental buscar equivalentes conceptuales en los estudios que utilizan varias comunidades o regiones.

Esto podría ser relativamente fácil cuando se trata de productos materiales, para los que se cuenta con algunos parámetros físicos. Sin embargo, la equivalencia conceptual de medidas cuando se trata de actitudes y creencias suele ser mucho más complicada. En algunas partes de América Latina, por ejemplo, cierto conjunto de alimentos apropiados para el tratamiento de una enfermedad puede catalogarse como "caliente" y esta categoría puede contraponerse a otro conjunto de alimentos designados como "frío". Algunas comunidades vecinas, sin embargo, pueden tener algunos productos alimenticios que trasciendan la distinción de "caliente-frío", o tener otros grupos de alimentos que no tienen importancia en la primera comunidad. Las respuestas a preguntas sobre el tratamiento apropiado, basadas en la suposición de categorías equivalentes, viciarían las conclusiones aplicables a varias comunidades. La necesidad de medidas equivalentes debe determinarse para cada grupo testigo y de tratamiento.

Por lo general, donde existe una teoría de desarrollo social, esa teoría tiende a ser formulada en términos abstractos, demasiado imprecisos para sugerir las variables específicas a medir. Aun cuando el

acopio de datos sobre todas las variables de posible interés es prácticamente imposible, es aconsejable incluir observaciones suplementarias que permitan evaluar hipótesis alternativas acerca de los efectos del programa. Por ejemplo, un estudio diseñado para medir el impacto de un programa de planificación familiar sobre los intervalos entre los partos, debe tomar en cuenta los cambios correspondientes en cuanto a creencias religiosas, tabús relativos a las relaciones sexuales después del parto y migración temporal de los cónyuges.

Cuestiones como las anteriores presentan problemas prácticos para los evaluadores y para las entidades operacionales. También demuestran que los costos de las investigaciones evaluativas aumentan en los estudios y obligan a extensos trabajos para asegurar la equivalencia de las medidas de resultados.

Aunque en muchos casos se puede aplicar con éxito una medida uniforme a una situación transcultural, metodológicamente es peligroso suponer la comprensión uniforme de estas medidas en las poblaciones estudiadas, sin datos de apoyo. Obviamente, el costo que implica verificar la idoneidad de las medidas de resultados debe cotejarse con otras necesidades del programa.

Relaciones entre el personal del programa y el de evaluación

Un segundo punto de gran importancia es la relación entre el personal del programa y el de evaluación. Varios autores que tratan sobre evaluación (25, 26) han examinado las dificultades con que tropieza el personal de evaluación en el ambiente en que se desarrolla el programa. En los países en desarrollo este problema surge por lo menos en dos niveles.

El primer nivel se refiere a la posible parcialidad del investigador a quien se le pide evaluar un programa del cual forma parte. Idealmente, el evaluador debe tener conocimientos técnicos en el área del proyecto, y una familiaridad con el personal del programa que permita una comunicación sin límites y un conocimiento detallado del proceso de ejecución del programa. En los países en desarrollo donde se cuenta con pocos expertos técnicos, donde es menos probable el registro sistemático de las operaciones, y donde sólo el personal del programa puede hacer *in situ* un examen concienzudo de las operaciones diarias, las ventajas de utilizar los servicios de un evaluador que forme parte del personal del programa parecen manifiestas. Sin embargo, Gordon y Morse (26) afirman que dicha estrategia tiene sus riesgos. Con base en una pequeña encuesta de estudios de evaluación publicados en Estados Unidos, dichos autores encontraron que en un 58% de los estudios en los que el evaluador estaba vinculado a la entidad patrocinadora, el programa fue juzgado como éxito, porcentaje que sólo fue del 14% cuando el evaluador era externo. Si este sesgo puede generalizarse a las condiciones prevalentes en los

países en desarrollo, el uso de evaluadores vinculados al programa podría conducir a la perpetuación de programas sociales ineficaces y al desperdicio de recursos de por sí ya limitados.

Un segundo riesgo es que los evaluadores vinculados al programa muchas veces se ven involucrados en el éxito de éste a tal grado que ya no creen que el programa necesite ser evaluado (27). Su papel a menudo se convierte en el del encargado de llevar el registro del proyecto, de quien se espera que tome nota de las actividades del personal del programa (v.g., número de clases de alfabetización impartidas, kilos de semillas distribuidos, etc.).

Por otro lado, el evaluador no vinculado al programa puede ser considerado como una amenaza para las funciones y el empleo del personal. Los rigurosos procedimientos de investigación pueden parecer hostiles a las metas del programa, ya que ponen en tela de juicio la competencia y las convicciones del personal del programa. El mejoramiento de la calidad de la evaluación mediante el empleo de mejores técnicas metodológicas podría agravar este problema, al aumentar la probabilidad de que se formulen reservas o se saquen conclusiones negativas respecto de la efectividad del programa. Además, el empeño del evaluador en que se ejecute el proyecto en forma de experimento controlado (es decir, prestando servicios a grupos seleccionados al azar, y no a otros grupos, etc.) puede parecer incompatible con los propósitos generales de servicio del personal, especialmente si los recursos de la entidad patrocinadora son abundantes. El término "evaluación", por sí solo puede presentar problemas. En el mundo del servicio público, donde los nombramientos generalmente son de corta duración, puede existir un alto grado de inseguridad al no contarse con un cuerpo de funcionarios profesional. La evaluación por parte de un forastero podría aumentar esa inseguridad, a menos que se tenga un entendimiento claro de dicho término.

Se han sugerido varias estrategias para reducir las suspicacias ante una evaluación externa sin detrimento de la eficacia del personal de evaluación. Una de ellas es concebir la evaluación desde la perspectiva más amplia del desarrollo del programa y ofreciendo sugerencias para mejorarlo (v.g., mejorando los procedimientos para identificar las familias que necesitan servicios, o para fortalecer la intervención). Esto es compatible tanto con los intereses del personal del programa como del de evaluación, ya que se administra el tratamiento más intenso posible, y es coherente con las exigencias del diseño experimental si no entraña cambios de importancia en el transcurso del período de evaluación. Otra estrategia es diseñar el experimento de manera que, a medida que el programa cobra capacidad para dar cabida a más clientes, se asignen periódicamente más miembros del grupo testigo aleatorio al grupo de tratamiento (11). Esta estrategia concuerda con las metas de desarrollo del programa, y no reduce la fuerza de la evaluación para

detectar los efectos del programa, si el grupo testigo es lo suficientemente grande y los sujetos transferidos se eligen al azar. Reicken y Boruch (10) recomiendan que se reitere al personal del programa que no se le hará responsable del tamaño ni la dirección de los efectos de la intervención. El propósito del programa es determinar si la intervención ejerce algún efecto, y si vale la pena repetirla en otro lugar, más bien que prestar un servicio en el lugar mismo de la investigación. Por ejemplo, la función de un programa experimental integrado de salud no es simplemente prestar atención de salud en las comunidades bajo estudio, sino comprobar si dicha intervención ejerce algún impacto generalizable a una población más grande. Situando los riesgos reales de la investigación evaluativa en su perspectiva propia, Cain y Hollister (28) señalan que la mayoría de los programas de acción social son tan complejos por la variedad de insumos y la multiplicidad de objetivos que es improbable que se considere conveniente abandonarlos.

Generalmente, el objetivo de los proyectos de evaluación es reordenar las actividades de intervención, y sugerir cambios marginales en la escala total del programa. La colaboración de colegas pertenecientes a la entidad responsable, como participantes en el estudio de evaluación, puede ser sumamente útil, tanto para mejorar la calidad de la intervención, como para aumentar la utilidad de los resultados de la evaluación. Cualquiera que sea la estrategia utilizada, el consenso de la mayoría de los evaluadores es que la confianza y cooperación del personal del programa son factores indispensables para evaluar adecuadamente los programas sociales.

Un segundo punto de preocupación, en cuanto a las relaciones del personal, es la percepción que la comunidad tenga de las funciones y actividades del personal del programa y del de evaluación. En los países en desarrollo la mayoría de los programas de cambio planificado se llevan a cabo a través de organismos gubernamentales. El sólo hecho de que un programa esté siendo auspiciado por el gobierno puede hacer que algunos individuos se muestren renuentes a participar. Una fuente de suspicacia ante los motivos de los proyectos patrocinados por el gobierno son las demandas de impuestos, conscripción militar y afiliaciones políticas de los representantes gubernamentales. En las áreas rurales de los países en desarrollo, el trabajo del personal del programa puede verse entorpecido no sólo por el hecho de que probablemente son empleados del gobierno, sino también por que son forasteros cuyos motivos y capacidades se desconocen. Los especialistas en desarrollo, cualquiera que sea la organización a la que estén vinculados, inspiran sospechas cuando entran en una comunidad, y puede llevar algún tiempo hasta que ésta se convenza de su interés sincero y genuino por el bienestar de la comunidad. Por ejemplo, en un proyecto de salud rural realizado en una aldea tradicional de las tierras altas de Guatemala, la desconfianza hacia los forasteros se

manifestó en los precios excesivamente altos que se pedían por un terreno donde sería instalada la clínica, así como en las reiteradas demandas de contar con pruebas inmediatas y palpables de las metas del programa (por ejemplo, suministros de medicinas y una ambulancia) en la etapa de organización comunitaria del proyecto.

Así, un problema importante del personal del programa es establecer su función como asesor visitante, abogando por comunidades que tradicionalmente han conocido pocas o ninguna de estas funciones (19). Un problema afín del personal de evaluación es el de distinguirse del personal del programa sin perder conexión con la comunidad. Por ejemplo, la identificación del personal de evaluación con el equipo de acción puede acarrear dificultades si los miembros de la comunidad se muestran renuentes a expresar actitudes negativas hacia el proyecto, por cortesía o por temor de que una evaluación negativa contribuya a la supresión del proyecto, en vez de mejorarlo. Con el tiempo, sin embargo, las distintas funciones se comprenden poco a poco, si el personal presenta una imagen coherente y el programa continúa a pesar de las críticas de la comunidad.

Financiamiento de los programas de investigación evaluativa

El costo del mantenimiento de los programas sociales es enorme y representa una carga financiera importante para los países en desarrollo que operan con presupuestos restringidos. Una manera de maximizar el efecto de estos fondos limitados en el desarrollo de servicios es el de comparar sistemáticamente métodos alternativos para el cumplimiento de las metas de desarrollo. Por ejemplo, los encargados de formular las políticas pueden cotejar los beneficios de la inversión en proyectos de nutrición con los de la inversión en servicios de extensión agrícola en lo que respecta a reducir la incidencia de la malnutrición. Tradicionalmente, la investigación evaluativa no se ha preocupado de los efectos relativos de tales alternativas dentro de una sola población.

Hasta hace poco y debido a la preocupación por detectar diferencias entre los grupos de intervención y los grupos testigo, la cuestión de si, considerando los costos, las diferencias de resultados son socialmente significativas no ha recibido particular atención (29). Por ejemplo, un programa masivo de suplementación nutricional puede reducir significativamente la malnutrición clínica en una población, pero el costo del suplemento, de su preparación y de su distribución pueden hacer que se descarte ese programa como estrategia de salud pública. Cuanto menos cueste alcanzar un objetivo del desarrollo en particular, más contribuirá el programa a la obtención de las metas de organismos con escasos recursos. En tal virtud, los programas sociales prototípicos en los países en desarrollo deben ser diseñados para reducir los costos,

dado el impacto deseado, o bien para obtener ingresos para su funcionamiento en caso de que se adopten, mediante, por ejemplo, impuestos progresivos sobre una mayor producción agrícola (especialmente con unidades de producción más extensas y eficientes) o ventas de alimentos elaborados hechas posibles gracias a una mayor producción.

El alto costo de la investigación evaluativa es también un problema de gran importancia. Perloff, Perloff y Sussna (30) se oponen a la norma frecuentemente formulada de que el 1% de los fondos de un programa debe asignarse para la evaluación del mismo. Hacen notar que ese 1% puede ser satisfactorio para una investigación programática en gran escala, pero que, tratándose de proyectos modestos (por ejemplo, aquéllos cuyo costo no llegue a 1 millón de dólares), supondría dedicar menos de US\$ 10 000 a la evaluación. Tal cantidad apenas bastaría para efectuar una evaluación global de intervenciones de bajo costo en los países en desarrollo, ya que requiere personal profesional multidisciplinario que cuente con apoyo administrativo y técnico adecuado.

Más que un porcentaje fijo, dichos autores recomiendan una cantidad negociable entre el patrocinador del programa y el evaluador, basada en la suma de dinero que sería razonable para ejecutar un plan de evaluación capaz de rendir al menos resultados mínimamente aceptables. El responsable de formular la política siempre se enfrenta con el dilema de correr ciertos riesgos si opta por obtener datos de la evaluación con insuficientes fondos y, por lo tanto, desperdiciar fondos en programas inútiles o, lo que es peor aún, correr el riesgo de obtener resultados no confiables, continuar programas ineficaces o apoyar otros programas que serán también un fracaso (31).

Los países en desarrollo pueden estar interesados en técnicas para reducir el costo relativamente alto de la investigación evaluativa, aun cuando cada técnica presente además problemas potenciales. Primero, los esfuerzos de evaluación pueden estar vinculados a un programa gubernamental ya existente o modificado. Esto puede reducir el costo inicial de la elaboración de un modelo de intervención, y permite también poner a prueba el efecto de ésta en un medio comparable a otros posibles lugares de adopción. Los inconvenientes de esta estrategia son la alta probabilidad de depender de diseños de investigación basados en datos posteriores a la ejecución del programa y la falta de control de la intervención. Para citar un ejemplo, la evaluación de un proyecto hidráulico gubernamental en Guatemala se vio obstaculizada por la falta de la oportunidad de reunir datos antes de la prueba y porque el gobierno no proporcionó materiales para el mantenimiento del sistema.

Una segunda técnica para reducir los costos es que los participantes reúnan sus propios datos. Este método no sólo libera los fondos que normalmente se utilizan para contratar personal de evaluación, sino que también da participación al personal del programa y a

la comunidad en el proceso de desarrollo del mismo. Esta técnica ha sido utilizada en Panamá en la evaluación de programas nutricionales de la comunidad, en un programa de salud rural en Guatemala (32) y en programas de desarrollo rural en África y América Latina auspiciados por la AID de Estados Unidos de América (33). En numerosos proyectos de vigilancia nutricional los participantes han reunido datos sobre indicadores de desnutrición, utilización de servicios comunitarios, patrones de cultivo, etc. En cambio, los problemas asociados a la utilización del personal del programa y sus usuarios para el acopio de datos hacen probable que sean poco fidedignas las técnicas y procedimientos de medición, lo que agrega un elemento de incertidumbre a la estimación de los efectos. La competencia entre las actividades asistenciales y las de acopio de datos resta a menudo prioridad a la evaluación. Además, puede existir predisposición consciente o inconsciente en favor del éxito del programa. A pesar de esas posibles distorsiones y problemas, estas técnicas pueden ser útiles si se lleva un control cuidadoso y se aplican precauciones para evitar los sesgos antes mencionados.

Un tercer método para reducir el costo de la investigación evaluativa en los países en desarrollo es emprender más estudios exploratorios a corto plazo, reseñas de publicaciones, investigación de archivos y ensayos piloto de posibles intervenciones, antes de hacer grandes inversiones en proyectos de evaluación acumulativa a largo plazo. Scriven (34) recomienda que, dada la relativa debilidad de muchos programas de intervención, quizás una inversión inteligente de los fondos destinados a evaluación sería utilizarlos en la fase formativa de los programas sociales. Si bien esta estrategia puede servir de ayuda a los directores de programas, ofrece a los responsables superiores de las decisiones poca información sobre los probables efectos del programa.

Una alternativa sería la de separar los fondos destinados al programa de intervención y al de evaluación. Podría argumentarse que, como los gobiernos de los países en desarrollo de todas maneras asignarían fondos para programas públicos, los fondos necesarios para la evaluación de estos programas podrían buscarse en el sector privado u obtenerse de organismos internacionales. Varios son los argumentos a favor de este enfoque. Primero, considerando que una evaluación global requiere una organización grande con numeroso personal multidisciplinario y un alto nivel de organización administrativa, sería difícil que cada país en desarrollo mantenga este costoso tipo de investigación sobre bases permanentes. Refuerza el argumento a favor de un financiamiento múltiple de evaluación el hecho de que un examen atento de los planes de ésta por los gobiernos y expertos nacionales, antes de asignar los fondos, generalmente mejoraría su calidad, y así reduciría el riesgo de que se produzcan resultados no válidos de la evaluación los cuales se utilizarían en la elaboración de políticas,

o de que resultados válidos se generalicen más allá de los límites legítimos.

Sin embargo, esta propuesta tiene varios inconvenientes. Primero, es poco probable que un organismo encargado de un programa, que no desea dedicar fondos ni personal a la evaluación, se comprometa a mantener el diseño de la investigación una vez que el proyecto se pone en marcha, o a actuar según los resultados. Segundo, las diferencias de valores entre el personal del programa y el personal evaluador aumentarán si la evaluación se considera esencialmente como una actividad dirigida y patrocinada por el exterior. Tercero, pocos organismos internacionales estarían dispuestos a adjudicar fondos si al mismo tiempo no se les garantiza cierto control o supervisión del contenido y ejecución del programa (por ejemplo, mantenimiento de sistemas de agua potable). Por último, puede ser que las fuentes de financiamiento internacionales o privadas de países desarrollados no sean políticamente aceptables para los gobiernos de los países en desarrollo.

Obviamente, no existe una solución ideal a estos problemas. Los evaluadores y el personal interesados en la ejecución de programas de índole social en países en vías de desarrollo deben encontrar un punto de equilibrio entre la demanda de fondos destinados a intervenciones y la necesidad de una evaluación encaminada a determinar si dichos programas son realmente eficientes.

Lo que se debe tener presente, en cuanto a técnicas para hacer economías en los estudios de evaluación, es que los verdaderos ahorros se lograrán en las operaciones futuras del programa. La mejor de estas técnicas puede ser, en última instancia, la documentación misma sobre el impacto de intervenciones repetibles y baratas, elaborada con buenos métodos de evaluación.

Consideraciones políticas y éticas asociadas con los estudios de evaluación en los países en desarrollo

A juzgar por la literatura publicada sobre este tema, es obvio que los problemas políticos y éticos que plantea la evaluación en los países en desarrollo son de gran importancia, puesto que ocupan casi el mismo espacio impreso que otros problemas más técnicos asociados con el diseño y el muestreo. Examinaremos sucesivamente cada uno de ellos.

Política y evaluación

Carol Weiss (35) asevera que la evaluación constituye una actividad racional que ocurre en un contexto político, y el evaluador que no llega a comprenderlo está expuesto a una serie de choques y frustraciones. Weiss identifica tres áreas principales en las que la política se inmiscuye en el proceso evaluativo o en las que la evaluación se inmiscuye en el proceso político.

Primero, los programas de acción social son producto del proceso político. Estos ya cuentan con sus defensores y antagonistas, y en la ejecución siguen siendo políticamente vulnerables. Segundo, la evaluación constituye un tema político, puesto que su primera meta racional es la de ayudar en la adopción de decisiones. La evaluación compete con otros factores del proceso político para reclamar atención. Finalmente, la evaluación es en sí una postura política. Examina críticamente las metas y estrategias de los programas y legitima la función de los investigadores sociales en la formación de políticas y programas.

Estas cuestiones se agravan en aquellas evaluaciones en las que, por motivos técnicos o por la necesidad de que no esté vinculado al programa, el investigador no es ciudadano del país donde aquel se desarrolla. En su propio país el evaluador es ciudadano además de científico, así pues, vota, paga impuestos, pertenece a grupos de interés y de presión y tiene derecho a participar en el proceso de acción social. En cambio, un evaluador extranjero no goza de ninguno de estos derechos. Sin embargo, si aceptamos las definiciones de Weiss y de otros autores (6, 26, 36), todo evaluador está involucrado en el proceso político de acción social. Le guste o no le guste, su papel es el de producir información que influirá en la política pública. Como los resultados de la evaluación se orientan hacia la toma de decisiones, y ésta es un proceso político y administrativo regido por el poder y el compromiso, el papel del evaluador puede ser muy complicado. Aquí es evidente que una definición razonablemente precisa de las metas del programa operativo es importante. Si las expectativas en cuanto al impacto del programa son exageradamente altas, el evaluador puede tropezar con serias dificultades. Caro señala que cuando las pretensiones administrativas para los programas de desarrollo son demasiado optimistas, los resultados de la evaluación son casi inevitablemente desalentadores (37). Weiss concuerda en que una de las muchas razones por las que los resultados de la evaluación se juzgan desalentadores es que los evaluadores han aceptado las promesas hinchadas y la retórica política como metas auténticas del programa (36). Dada la naturaleza compleja del subdesarrollo, los programas de intervención deben tener expectativas más modestas y evaluarse en comparación con metas más razonables. Tripodi manifiesta que sólo se obtiene información útil sobre programas sociales cuando el clima sociopolítico favorece la investigación honesta y existe la voluntad de utilizar la evaluación como un instrumento de gestión para ampliar los conocimientos y tomar decisiones (38).

Por lo tanto, el evaluador tiene que afrontar numerosos problemas de credibilidad: la de su propia experiencia en el desarrollo de la comunidad, la credibilidad de las metas de la política gubernamental y, por último, los problemas que acarrea relacionados con la organización y actividades del organismo encargado del programa de intervención (26, 35, 39).

Además, como miembro de una comunidad científica, el evaluador suele estar interesado en la publicación de los resultados de la evaluación. La cuestión de quién tiene derecho a publicar los datos—el gobierno, la entidad patrocinadora o el investigador—puede ser muy delicada. ¿Cuáles son las implicaciones de que un evaluador presente datos sobre los resultados de un programa social al público por el canal de publicaciones científicas? Dicha información puede ser políticamente delicada y hasta embarazosa para el gobierno huésped.

En estas circunstancias, incumbe al evaluador aclarar las políticas de publicación desde un principio y quizás como condición previa para emprender cualquier evaluación. Tripodi sugiere la conveniencia de que el derecho sobre los datos y la política para su publicación formen parte del contrato de evaluación, con lo cual se evitaría desde un principio todo malentendido (38).

Ética y evaluación

Los programas de desarrollo aportan algún servicio social o económico. Algunas veces responden a necesidades de evaluación en las comunidades, como se propugna en la literatura antropológica relativa al cambio social (14, 17, 40, 41). Con cierta frecuencia, las intervenciones son producto de programas de planificación nacional y son impuestas a la comunidad, sin tomar en cuenta muchas veces la percepción local de las necesidades. Aparte de esto, los evaluadores que tratan de sacar el mejor partido posible de una situación difícil, deben insistir en la máxima aproximación posible a un diseño de investigación experimental que, si es posible, incluya un tratamiento bien definido, grupos testigo y la selección aleatoria de los participantes en el programa.

En los países en desarrollo, la mayoría de los programas de acción social están dirigidos a los estratos socioeconómicos más pobres. Esto quiere decir que casi todos los que son objeto de la intervención tienen necesidad acuciante de los servicios que el programa proporcionará. El problema que entonces se presenta es ¿cómo justificar la no prestación de tales servicios a un grupo de individuos?

En contraste con las investigaciones médicas, los participantes en investigaciones sobre intervenciones sociales, por lo general, pueden fácilmente distinguir entre los placebos y el verdadero tratamiento. La preocupación por obtener "ganancias netas" para todos es genuina, pero difícil de reconocer, si de lo que se trata es de medir la eficacia de una intervención en particular. En algunos casos un experimento puede ensayar dos o más formas de intervención a fin de determinar cuál de ellos rinde los mejores resultados. En este caso todos los participantes podrán obtener algún beneficio. Sin embargo, no es frecuente que se hagan experimentos sobre diferentes formas de intervención, y ordinariamente se opta por ensayar una u otra de ellas. A este respecto, Freeman (42) lamenta

que muchas veces las presiones políticas son las que determinan las intervenciones adoptadas en programas sociales.

Otro problema de ética en las evaluaciones se refiere a la vida privada de los participantes en el programa y el carácter confidencial de los datos de evaluación. Muchas evaluaciones de programas giran en torno al individuo. A éste se le pide que divulgue aspectos de su vida personal a entrevistadores desconocidos y muchas veces impersonales. Puesto que una de las características comunes a los pobres en los países en desarrollo es la falta de educación, muchas veces se hace difícil, si no imposible, explicar las precauciones de discreción que se toman en el tratamiento electrónico de los datos. Aparte de esto, y según su experiencia anterior con organismos gubernamentales (v.g., el censo y la recaudación de impuestos), puede que el informante tenga razones para desconfiar de las promesas del evaluador en cuanto a la protección de los informantes.

Finalmente, el evaluador debe determinar quién es el cliente. ¿Es su cliente el gobierno y su organismo de ejecución los sujetos de la intervención, o la organización que aporta los fondos para la evaluación? Esta pregunta crucial nos hace volver al tema de la política comentado anteriormente. ¿Ante quién es responsable el evaluador en última instancia?

Muchas de las preguntas y cuestiones suscitadas en esta sección no han tenido una respuesta satisfactoria. Se hace necesario un concienzudo debate que incluya a representantes de todos los sectores afectados por el proceso de evaluación, a fin de determinar algunas directrices concretas. A través del debate continuo, el uso de los métodos de evaluación puede contribuir al éxito de los programas de acción social en los países en desarrollo.

Oportunidades y necesidades de los estudios de evaluación en los países en desarrollo

A pesar de los problemas técnicos, operacionales y éticos comentados anteriormente en este trabajo, existe consenso general en que las estrategias de la investigación evaluativa pueden contribuir en gran medida a los programas de desarrollo. Pueden, por ejemplo, estimular la especificación cuidadosa de las políticas y metas del programa y mejorar las estrategias de intervención gracias a la información aportada por la medición de las actividades y efectos del programa. Los efectos benéficos generales de la evaluación en un programa pueden apreciarse a todos los niveles administrativos en una planificación mejorada, en el registro sistemático de control de las actividades del mismo, y en la actitud sana y constructivamente crítica ante el programa manifestada por el personal responsable.

Existen tres factores generales que en la actualidad limitan la aplicación más amplia de las técnicas de investigación evaluativa a las estrategias de desarrollo.

El primero es la disponibilidad de personal capacitado. Hasta hace cinco años no comenzaron a surgir en las universidades norteamericanas programas específicamente diseñados para formar especialistas en investigación evaluativa. La mayoría de éstos provienen de países desarrollados; por lo tanto, la disponibilidad de este personal en los países en desarrollo es muy reducida. La capacitación y organización de evaluadores procedentes de países en desarrollo obviamente constituye un requisito previo para la expansión de estas actividades en el contexto de programas óptimos de desarrollo.

Un segundo factor que limita la aplicación más extensa de técnicas de investigación evaluativa en los programas de desarrollo, es la falta relativa de comprensión y apreciación de los usos y ventajas de las estrategias de evaluación por parte de los políticos y personal administrativo superior en esos países. Aun cuando la evaluación se está expandiendo en estos países, principalmente se ha aplicado en investigaciones limitadas a innovaciones agrícolas, planificación familiar y algunas intervenciones de salud y nutrición. La transferencia y adopción de esta tecnología en diversidad de programas no agrícolas ha sido lenta.

A pesar de que en los últimos años se han producido positivos e importantes cambios, todavía es necesario informar al personal directivo acerca del papel que desempeña la evaluación y diseminar ampliamente los resultados de que se dispone para dar a conocer los beneficios que podría acarrear la aplicación sistemática de técnicas de evaluación de programas.

Un tercer factor, obviamente relacionado con los dos primeros, es la necesidad de mejorar el diseño, los métodos, las medidas de los resultados y las técnicas analíticas en la investigación evaluativa. Así, se logrará progresar a medida que se obtenga más experiencia para determinar las necesidades especiales de la evaluación del desarrollo e investigadores más competentes participen en la evaluación de diversos componentes de los programas de desarrollo nacional. Los progresos en estos aspectos técnicos, y especialmente en el desarrollo y uso de medidas de resultados comparables o uniformes, permitirán comparar los resultados de las distintas intervenciones y sistemas de ejecución en cuanto a su impacto y costo relativo. Más adelante se dispondrá de un conjunto coherente de publicaciones, aprovechando las experiencias de programas de vasto alcance en los países desarrollados, que ayudará mucho a explicar los problemas metodológicos y operacionales de programas de menor escala apropiados para satisfacer las necesidades de los países en desarrollo. Esto ayudará a los diseñadores y administradores de programas a seleccionar estrategias de desarrollo y a diseñarlos sobre una base más racional que en la actualidad.

Como indicamos en la introducción de este capítulo, cada vez se da más importancia a la investigación evaluativa como instrumento útil en la estrategia de desarrollo. Esta tendencia podría acelerarse mediante el

establecimiento de una red coordinada de investigación y capacitación en América Latina. Esta red podría introducirse inicialmente en universidades y centros de investigación de varios países latinoamericanos.

Las actividades principales de estos centros serían la capacitación en diseño y ejecución de actividades de evaluación. Sin embargo, esos centros también podrían servir para el intercambio de nuevos datos de evaluación y el análisis secundario de los datos existentes, y como fuentes de referencia para los administradores de programas que deseen hacer consultas sobre evaluación.

Las actividades de esa red, por su propia naturaleza, estarían dirigidas hacia los tres factores principales que limitan el uso más amplio de la investigación evaluativa, con el fin de promover y expandir la aplicación

de las actividades de evaluación en el área general del desarrollo económico y social.

Nota de agradecimiento

Para la preparación de este trabajo se contó con apoyo de la Agencia para el Desarrollo Internacional, Washington, D.C. (Contrato AID-TA-C/1224); del Instituto Nacional de Salud del Niño y Desarrollo Humano, Bethesda, Maryland (Contrato No. NO1/DH-5-0640) y de la Fundación Ford, Nueva York (Subvención No. PN-801).

Las opiniones aquí expresadas son de los autores y no reflejan necesariamente las de las instituciones patrocinadoras, ni las de la Organización Panamericana de la Salud.

REFERENCIAS

- Bernstein, I. y H. Freeman. *Academic and Entrepreneurial Research*. New York, Russell Sage Foundation, 1975.
- Newhouse, J. P. *A Summary of the Experimental Portion of the Rand Health Insurance Study*. Edición revisada. Santa Monica, California, Rand Corporation, 1977.
- Watts, H. y A. Rees (Eds.). *Final Report on the Graduate Work Incentive Experiment in New Jersey and Pennsylvania*. Madison, University of Wisconsin, Institute for Research on Poverty, 1974.
- Freeman, H. E. The present status of evaluation research. En: *Proceeding of UNESCO Evaluation Research Conference*. Washington, D.C., 1976.
- Batten, T. R. *Communities and their Development: An Introductory Study with Special Reference to the Tropics*. London, Oxford University Press, 1957.
- Campbell, D. Reform as experiments. En: Struening, E. L. y M. Guttentag (Eds.). *Handbook of Evaluation Research*. Beverly Hills, California, Sage Publications, 1975.
- Manoff, R. K. y T. M. Cooke. *Changing Nutrition and Health Behaviors through the Mass Media: Nicaragua and Philippines*. New York, Manoff International, 1976.
- Kershaw, D. Administrative issues in income maintenance experimentation: Administering experiments. En: Orr, L. L., R. G. Hollister, M. Lefcowitz y K. Hester (Eds.). *Income Maintenance: Interdisciplinary Approaches to Research*. Chicago, Marham, 1971.
- Campbell, D. y J. Stanley. *Experimental and Quasi-Experimental Designs for Research*. Chicago, Rand-McNally, 1963.
- Riecken, H. W. y R. F. Boruch. *Social Experimentation: A Method for Planning and Evaluating Social Intervention*. New York, Academic Press, 1974.
- Boruch, R. Coupling randomized experiments and approximations to experiments in social program evaluation. En: I. N. Bernstein (Ed.). *Validity Issues in Evaluative Research*. Beverly Hills, California, Sage Publications, 1975.
- Bracht, G. H. y G. V. Glass. The external validity of experiments. *Am Educ Res J* 5:437-474, 1968.
- Foster, G. M. *Traditional Societies and Technological Change*. New York, Harper & Row, 1973.
- Spicer, E. H. *Human Problems in Technological Change*. New York, John Wiley & Sons, 1952.
- Harrison, R. P. Non-verbal communication. En: Pool, I. de S. y W. Schram (Eds.). *Handbook of Communication*. Chicago, Rand-McNally, 1973.
- Hall, E. T. *The Hidden Dimension*. Garden City, New York, Doubleday, 1966.
- Foster, G. M. *Traditional Cultures and the Impact of Technological Change*. New York, Harper & Row, 1962.
- Osgood, C. Semantic differential techniques in the comparative study of cultures. En: Rommey, A. K. y R. D'Andrade (Eds.). *Transcultural studies in cognition*. *Am Anthropol* 66: Specific Issue 2, 1964.
- Rommey, A. K., R. N. Shepard y S. B. Nerlove (Eds.). *Multidimensional Scaling. Theory and Applications in the Behavioral Sciences*, vol. 2. New York, Seminar Press, 1972.
- Pelto, P. *Anthropological Research: The Structure of Inquiry*. New York, Harper & Row, 1970.
- Belcher, J. C. A cross-cultural household level-of-living scale. *Rural Sociol* 37(2):208-220, 1972.
- Farrel, W. T. *Community Development and Individual Modernization in San Lucas Toliman, Guatemala*. Ph. D. Dissertation. Los Angeles, University of California at Los Angeles, 1977.
- Klein, R. E., M. Irwin, P. L. Engle y C. Yarbrough. Malnutrition and mental development in rural Guatemala: An applied cross-cultural research study. En: Warren N. (Ed.). *Advances in Cross-Cultural Psychology*. New York, Academic Press, 1977. Págs. 91-119.
- Rodman, H. y R. Kolodny. Organizational strains in the researcher-practitioner relationship. En: Caro, F. G. (Ed.). *Readings in Evaluation Research*. New York, Russell Sage Foundation, 1971.

Evaluación del impacto de los programas de nutrición y de salud

25. Suchman, E. A. *Evaluative Research*. New York, Russell Sage Foundation, 1967.
 26. Gordon, G. y E. V. Morse. Evaluation research. En: *Annual Review of Sociology*, vol. I. Palo Alto, California, Annual Reviews, 1975. Págs. 339-361.
 27. Weiss, R. A. y M. Rein. Trabajo presentado en la conferencia The Evaluation of Social Action Programs auspiciada por the American Academy of Arts and Sciences, Washington, D. C., 1969.
 28. Cain, G. G. y R. G. Hollister. The methodology of evaluating social action programs. En: Rossi P. H. y W. Williams (Eds.). *Evaluating Social Programs*. New York, Seminar Press, 1972.
 29. Levin, H. M. Cost-effectiveness analysis in evaluation research. En: Guttentag, M. y E. L. Struening (Eds.). *Handbook of Evaluation on Research*. Beverly Hills, California, Sage Publications, 1975.
 30. Perloff, R., E. Perloff y E. Sussna. Program evaluation. En: Rosenzweig, M. R. y L. W. Porter (Eds.). *Annual Review of Psychology*, vol. 27. Palo Alto, California, Annual Reviews, 1976. Págs. 569-594.
 31. Rossi, P. H. y W. Williams. *Evaluating Social Programs*. New York, Seminar Press, 1972.
 32. Beghin, I. Improving nutrition at the local level. *UNICEF Carnets Enfance* 35, 1976.
 33. Morse, E. R., J. K. Hatch, D. R. Mickelwait y C. F. Sweet. *Strategies for Small Farmer Development*. Boulder, Colorado, Westview Press, 1976.
 34. Scriven, M. The methodology of evaluation. En: Tyler, R. W., R. M. Gagne y M. Scriven (Eds.). *Area Monograph Series on Curriculum Evaluation*, 1. Chicago, Rand-McNally, 1967.
 35. Weiss, C. H. Evaluation research in the political context. En: Struening E. L. y M. Guttentag (Eds.). *Handbook of Evaluation Research*. Beverly Hills, California, Sage Publications, 1975.
 36. Rossi, P. Evaluating educational programs. En: Caro, F. G. (Ed.). *Readings in Evaluation Research*. New York, Russell Sage Foundation, 1971. Págs. 97-99.
 37. Caro, F. G. Evaluation research; an overview. En: Caro, F. G. (Ed.). *Readings in Evaluation Research*. New York, Russell Sage Foundation, 1971. Págs. 1-34.
 38. Tripodi, T., P. Fellin e I. Epstein. *Social Program Evaluation*. Ithaca. New York, Peacock Publishers, 1971.
 39. Davis, H. E. y S. Sulasin. The utilization of evaluation. En: Struening E. L. y M. Guttentag (Eds.). *Handbook of Evaluation Research*. Beverly Hills, California, Sage Publications, 1975.
 40. Goodenough, W. H. *Cooperation in Change*. New York, Russell Sage Foundation, 1963.
 41. Arensberg, C. y A. Niehoff. *Introducing Social Change*. Chicago, Aldine, 1971.
 42. Freeman, H. E. Foreword. En: Caro, F. G. (Ed.). *Reading's in Evaluation Research*. New York, Russell Sage Foundation, 1971. Págs. ix-xi.
-