

EVALUACION DEL MODELO

Robert E. Klein¹

Introducción

Según pudimos apreciar en el trabajo de Urrutia y García, los factores contribuyentes al subdesarrollo en Santa María Cauqué son múltiples e interrelacionados. Por ejemplo, se observa que la desnutrición resulta de una interacción de factores tales como incremento de población, disminución de la cantidad de tierra laborable por familia, reducción del agua disponible para riegos, falta de saneamiento ambiental adecuado, analfabetismo, carencia de conocimientos en la aplicación de técnicas agrícolas mejoradas, etc.

Esta situación se repite sistemáticamente en todos los países subdesarrollados del mundo. Con miras a resolver esta grave situación, recientemente se ha hecho más énfasis en los programas de desarrollo integral, ya que los gobiernos y agencias privadas han reconocido la complejidad del problema. En general, esto implica no sólo intervenciones múltiples en un sector único sino proyectos coordinados en varios sectores. La razón fundamental de esta política es lógica y muy directa. Por ejemplo, es probable que la educación nutricional a nivel de la familia rural tenga poco efecto si su situación económica no les permite aplicar los conocimientos impartidos. De manera semejante, tratar de mejorar las tasas de analfabetismo parece tarea inútil si saber leer y escribir no tiene un fin pragmático. En otros términos, muy poca gente pobre del medio rural invertirá su tiempo en un proyecto aislado carente de una finalidad que será de beneficio para ellos si lo pueden comprender. Las implicaciones de esta afirmación son dos: primera, una intervención en un sólo sector no resolverá el problema global; y segundo, los proyectos aislados sin un fin pragmático inmediato tienen baja probabilidad de éxito. Por lo tanto, los programas integrales como el propuesto por Urrutia y García, parecen más relevantes y tienen mayor potencial de contribuir a la solución del problema del subdesarrollo rural.

El Papel de la Evaluación

La evaluación es una actividad administrativa y científica que se emplea para determinar la eficacia y el impacto de cualquier tipo de intervención que pretende

¹ Jefe de la División de Desarrollo Humano, Instituto de Nutrición de Centro América y Panamá (INCAP), Apartado Postal 1188, Guatemala, Guatemala, C. A.

el logro de cierto cambio. La mayoría de los programas de desarrollo tienen objetivos múltiples aun cuando han sido planificados y ejecutados por un sólo sector. Obviamente, cuando se trata de programas que involucran más de un sector, la complejidad de los objetivos se ve grandemente incrementada. Es natural que los políticos, planificadores y administradores de programas de desarrollo quieran saber si el programa funciona o no, y el por qué. Esta pregunta global puede dividirse en cinco subpreguntas que se presentan seguidamente:

1. ¿Coinciden los programas con las metas políticas?
2. ¿Están ejecutándose los programas eficiente y correctamente?
3. ¿Tienen el impacto esperado?
4. ¿Hay consecuencias secundarias, no previstas, que se relacionen con el programa?
5. ¿Cuáles son las implicaciones del éxito o el fracaso de un programa para tomar decisiones futuras respecto a programas de esta naturaleza?

Estas son preguntas fundamentales en las que deben enfocarse las investigaciones evaluadas. El objetivo de la investigación evaluativa es incorporar estas preguntas a un programa, y proporcionar respuestas. Las primeras cuatro preguntas forman parte directa de la investigación evaluativa. La quinta pregunta depende de las interpretaciones que los dirigentes den a las conclusiones del evaluador en relación con las políticas generales del organismo patrocinador.

La evaluación comprensiva de un programa tiene dos componentes: la evaluación del proceso y la evaluación del impacto.

La evaluación del proceso se refiere a las preguntas No. 1 y No. 2 mencionadas, o sea, la ejecución del programa. Por ejemplo, ¿coinciden las metas a nivel operacional con las metas del nivel político? ¿Es que las actividades y/o materiales llegan a su destino a tiempo? ¿Cumple con sus responsabilidades el personal del programa?, etc.

La evaluación del impacto, en cambio, mide hasta qué punto cumplió el programa los objetivos planificados, y qué otros efectos ha tenido el resultado de la intervención (preguntas No. 3 y No. 4). La combinación de la evaluación del proceso y la evaluación del impacto permite al evaluador y al administrador formular conclusiones respecto al programa en sí, y aportar la información necesaria para la toma de decisiones a nivel político en cuanto al futuro de programas semejantes.

En algunos modelos de evaluación se necesita conocer el valor relativo de varias de las intervenciones, a fin de determinar qué tipo de intervención tiene el mayor efecto o impacto. En el caso de Santa María Cauque (SMC), por ejemplo, se pretende utilizar varias estrategias combinadas con un componente fuerte de participación comunitaria para alcanzar, como objetivo general, que las familias de SMC "disfruten

de un bienestar que les permita una productividad económica y social elevada". En el supuesto que este objetivo general pueda definirse en términos específicos y cuantitativos, el problema estriba en determinar cuál o cuáles de las intervenciones son las que alcanzan los resultados esperados. En la Tabla 1 se expone esta situación. Según se aprecia, no es factible determinar cuál (o cuáles) de las intervenciones tiene el efecto más fuerte en cuanto a la obtención de los resultados, porque existe entre ellas interacción al nivel de la familia. Por ejemplo, una familia puede participar activamente en todas las intervenciones y lograr el bienestar que se espera, pero ¿cómo decidir cuál de las intervenciones fue más importante y generalizar la estrategia a programas nacionales? La importancia de esto último es que las naciones en vías de desarrollo cuentan con fondos limitados, y un programa que se pretende generalizar tiene que medir costos y beneficios. Para llegar a esta decisión se necesita de un diseño factorial, como el que se presenta en la Tabla 2. Aquí se nota que se necesitan ocho comunidades para medir los efectos de las intervenciones de cada una o de las diversas combinaciones. Este proceso es costoso y asume también una serie de controles tales como la existencia de homogeneidad entre las ocho comunidades y la ausencia de otras intervenciones por parte de otros grupos, etc. Como consecuencia, un modelo evaluativo experimental no es apropiado para el proyecto propuesto para Santa María Cauqué.

TABLA 1

Esquema de intervenciones y resultados de un programa de desarrollo integral

<i>Intervenciones</i>	<i>Resultados Esperados</i>
<p>Agropecuaria</p> <ol style="list-style-type: none"> 1. Técnica y producción 2. Transformación 3. Mercadeo 4. Consumo <p>Atención médica</p> <ol style="list-style-type: none"> 1. Promoción 2. Técnica 3. Educación <p>Saneamiento ambiental</p>	<p>Bienestar</p> <p>Definida en sub objetivos como:</p> <ul style="list-style-type: none"> — Estado nutricional* — Estado de salud* — Aumento del ingreso familiar* — Vivienda mejorada* — Otros*
<p>Participación comunitaria</p>	

* Todos los objetivos deben definirse en términos cuantitativos.

En el caso de la evaluación de un programa de desarrollo en una sola comunidad, dos son las técnicas principales disponibles. La primera trata de mantener un

cierto control experimental a través del proceso de selección al azar de los participantes en el proyecto. En otras palabras, los individuos o grupos hogareños son seleccionados al azar para que participen en el programa. Aquéllos no asignados se convierten en el grupo testigo. Aun cuando esto proporciona cierto grado de control para el estudio del programa, también crea problemas éticos y prácticos. En las comunidades donde la mayoría de la población es muy pobre es bastante difícil, si no imposible, justificar la retención de los beneficios potenciales de un programa a un segmento de la población, hasta utilizando los medios más imparciales posibles. Aunque este curso de acción fuese éticamente justificable, todos aquéllos que quedasen excluidos podrían causar problemas prácticos interfiriendo con la ejecución de los proyectos del programa.

La solución a este dilema parece dejar una sola y débil alternativa desde el punto de vista científico. Esta es que el individuo o familia se autoseleccione para participar en el programa; la misma familia o el individuo se usan así como su propio control, antes y después de las intervenciones. En casos de esta naturaleza, el especificar la relación entre las intervenciones y los indicadores de impacto se hace más difícil. Esto es especialmente cierto cuando se trata de intervenciones múltiples.

Dos son los problemas básicos que afectan los programas autoseleccionados en una sola comunidad, y están relacionados a la validez externa e interna, respectivamente.

La validez externa, o sea el grado en que los hallazgos pueden ser generalizados o duplicados en otras poblaciones, es un problema serio en el diseño que contempla una sola comunidad. Esencialmente no se puede hacer nada que garantice que los resultados obtenidos en una sola comunidad puedan obtenerse mediante intervenciones similares en otras poblaciones. Sin embargo, aunque éste es el caso, existe un cúmulo substancial de literatura en ciencias sociales que ofrece algunos principios generales para el desarrollo comunitario capaz de aumentar la probabilidad de generalizar los hallazgos de los estudios realizados en una sola comunidad.

La validez interna —en este caso el grado al cual se miden los efectos del programa y no los efectos de otra fuente— también presenta un problema serio con la autoselección. Es decir, ¿cuáles son las características que influyen en una persona o familia para que participe en un programa, mientras que otras familias se niegan a participar? En la Tabla 3 se enumeran algunas características que pueden afectar la autoselección. Obviamente, la lista no es exhaustiva, y todas las variables están relacionadas entre sí en varias formas.

Estas variables deben ser estadísticamente controladas en el diseño evaluativo de una sola comunidad, porque se trata de comparar un individuo o grupo familiar consigo mismo, a través de cierto tiempo. Además, las diferencias en estos factores posiblemente determinen la manera en que las personas se comportan con respecto al programa.

TABLA 3

Problemas con evaluaciones de autoselección en una sola comunidad

A. Factores históricos

1. Educación
2. Influencia de los padres
3. Efectos (exposición) a influencias externas

B. Factores biomédicos

1. Estado de salud
2. Edad

C. Factores de situación (estado) social

1. Situación económica
2. Status social
3. Relaciones sociales
4. Afiliación religiosa
5. Afiliación política

D. Otras variables (psicológicas)

1. Motivación hacia el cambio
 2. Curiosidad
 3. Descontento con su papel social presente
-

Medición de Intervención e Impacto

Puesto que la unidad de análisis en una evaluación es el individuo o el grupo familiar, la medida de las intervenciones y su impacto se hace a este nivel. La autoselección implica que un individuo pueda participar en uno, varios o todos los proyectos de un programa, o bien no participar en ninguno. Puede ingresar a cualquier proyecto o egresar del mismo en un momento dado; aun puede convertirse en un opositor activo de cierto proyecto. En consecuencia, deberá aplicarse un método sistemático y detallado para registrar el *grado de participación*. Esta es la única forma de saber *qué y cuánto* del programa recibió la unidad objeto de análisis.

El sistema de registro a que se alude debe incluir medidas apropiadas para cada proyecto en particular. Tomemos como ejemplo la intervención agrícola que propuso el Ing. Jarquín, que consistirá en técnicas para la siembra de maíz. El plan de aplicación deberá constar de: 1) sesiones con los agricultores y asistencia se mide por el número de sesiones a que acude cada individuo; 2) el conocimiento, por parte

del agricultor, de las propiedades de la nueva semilla (susceptible de medir a través de discusiones individuales con los extensionistas); 3) la siembra en sí usando la nueva técnica (que se mide según la persona que lo sembró); 4) mantenimiento apropiado de terreno (que se mide según observaciones del extensionista); y 5) producción (medida según la cantidad cosechada). Hasta aquí, hemos medido el proceso de intervención y algunos impactos primarios. No obstante, también es necesario observar, a manera de contraste, las estrategias empleadas por los no-participantes.

Este es solamente el primer paso. Si la cosecha es superior a la obtenida con el método tradicional, debemos saber cómo se emplean los ingresos monetarios adicionales, para lo cual existe cierto número de alternativas que varían según el individuo o grupo familiar. Algunas de éstas son inversiones de capital, mejoramiento de vivienda, compra de mayor cantidad y mejor calidad de alimentos y ropa, artículos de consumo, mantenimiento escolar de los niños, etc. Estos pueden ser llamados conceptos variables de impacto secundario.

Al mismo tiempo, queremos saber si la esposa del agricultor asiste a clases de salud y nutrición, su comprensión del tema presentado y, finalmente, si estas lecciones se aplican en el hogar, como mencionó en su trabajo la Lic. de Ibarra.

Lo mismo aplica a cada uno de los proyectos iniciados en la comunidad. Este es un proceso largo y meticuloso que se mantiene durante todo el programa.

Dado que los programas se diseñan para aumentar la autosuficiencia del individuo y de la comunidad, los planificadores y gestores de programas deberán calcular de manera realista el tiempo necesario para establecer el programa y el tiempo que deberán mantenerse las intervenciones directas. También tendrán que especificar anticipadamente cuáles son los impactos finales previstos, de preferencia en términos cuantitativos. Al discontinuar las intervenciones directas, el evaluador debe comenzar entonces a analizar los datos recabados.

La síntesis de tales datos requiere que el evaluador sea capaz de cuantificar los indicadores brutos del impacto: ingresos familiares, producción agrícola, estado nutricional y de salud, vivienda y migración. Estos resultados pueden entonces compararse con los previstos, y determinarse así el éxito o fracaso relativo. No obstante, también es posible avanzar un paso más, por medio del análisis cuidadoso de los datos en cada individuo ó familia. Usando los datos de autoselección antes mencionados, se podrá clasificar a los grupos familiares o individuos, y las tipologías de esta índole pueden ser útiles para determinar el impacto relativo, por tipo y cantidad de la intervención, según el tipo de individuo. De esta manera, el evaluador puede especificar algunas de las relaciones entre intervenciones e impacto con respecto a grupos específicos de personas.

Un ejemplo de este método se presenta en la evaluación de una sola comunidad hecha en el altiplano de Guatemala. Se recabaron características personales en una muestra de participantes y no-participantes al hacer el estudio (1972) y, retrospectivamente (1960), antes de que comenzara el programa. La Tabla 4 revela los datos retrospectivos para 1960. Según se aprecia, los datos indican pocas diferencias de

TABLA 4

Pruebas de significancia de las diferencias en el estado económico circa 1960
entre 19 participantes y 19 no-participantes

Variable	Prueba y estadística	Nivel de significación
Salario diario	t = 1.24	N.S.
Terrenos de milpa propios	t = 1.16	N.S.
Terrenos de café propios	t = 1.3	N.S.
Propietario del terreno de la casa	X ² = 1.17	N.S.
Radio propio	X ² = 2.0	N.S.
Reloj de pulsera/reloj de mesa propios	X ² = 1.58	N.S.
Usó fertilizantes químicos	X ² = 0	N.S.
Clase de techo de la vivienda (paja o lámina)	X ² = 3.0	0.07
Clase de paredes (caña u otra)	X ² = 5.8	0.03
Agua corriente en el hogar	X ² = 3.2	0.06
Letrina en la vivienda	X ² = 1.8	N.S.
El padre fue dueño de terrenos de café antes de 1960	X ² = 1.13	N.S.

cierta significancia entre los participantes antes de iniciarse la intervención.

De manera semejante, la Tabla 5 tampoco muestra diferencias en las variables de antecedentes entre los participantes y no-participantes, antes de su *autoselección*, en los diferentes proyectos en este programa de desarrollo en particular.

Como se observa en la Tabla 6, sin embargo, en 1972 este mismo grupo acusó diferencias significativas en varias dimensiones. Ya que no es posible llegar a deducciones a partir de los análisis causales absolutos, podemos decir que claramente existe una diferencia entre los dos grupos autoseleccionados.

Cuando se trata de un estudio de una sola comunidad, ésta es la clase de datos significativos que los que diseñan la política necesitan y desean obtener. El programa parece haber dado resultado.

A pesar de ello, no es posible especificar la manera exacta de cómo funcionó o cuáles de los insumos tuvieron más éxito en términos de cómo *causaron* estas diferencias, pero es probable que el proyecto haya sido la causa principal de las diferencias posteriores observadas en el grupo.

El programa en cuestión fue integrado sin estar debidamente reglamentado por buenas normas administrativas (de proceso). No obstante, el resultado fue una diferencia subjetiva y objetiva, y puede decirse que sí mejoró el nivel de vida de los

TABLA 5

Pruebas de "t" de las variables económicas y sociales que comparan a los participantes y no-participantes con anterioridad a la introducción del programa de desarrollo (1960)

Variable	Prueba y estadística	Nivel de significación
Ocupación del padre (trabajador agrícola vs. no-agricola o propietario de tierras)	$X^2 = 0.36$	N.S.
Herencia de tierras de café	$X^2 = 0.88$	N.S.
Herencia de tierras de maíz	$X^2 = 0.82$	N.S.
Tipo de ropa que usó de niño (tradicional-ladino)	$X^2 = 1.96$	N.S.
Era propietario de café cuando se casó	$X^2 = 0.03$	N.S.
Era propietario de maíz cuando se casó	$X^2 = 1.22$	N.S.
Salario diario cuando se casó	$t = 0.05$	N.S.
Edad a la que se mudó de la casa del padre	$t = 0.60$	N.S.
Edad a la que cambió a vestimenta moderna	$t = 0.32$	N.S.

participantes.

En el caso de Santa María Cauqué, donde abundan los datos longitudinales, esta estrategia de evaluación parece bastante factible. Lo que los diseñadores y técnicos del programa necesitan hacer es especificar lo más exactamente posible cuáles indicadores son los más representativos de lo que se considera el mejoramiento del bienestar de los residentes. Cabe mencionar, asimismo, que los indicadores generales como cantidad de terreno sembrado en hortalizas o número de terneros criados, no son suficientes para declarar que ha habido un impacto. Siempre existe la posibilidad de que solo un pequeño grupo participe y aproveche las intervenciones, resultando en un desbalance respecto a la mayoría de los habitantes, quienes quedan en una situación peor de la que tenían antes de la intervención. Los indicadores de esta naturaleza son apropiados para uso a nivel regional o nacional. Sin embargo, al nivel de una sola comunidad, los indicadores de hecho tienen que reflejar el estado de bienestar al nivel familiar o individual.

TABLA 6

Pruebas de significancia entre participantes (n = 43) y
bajos / no-participantes (n = 42) en las dimensiones económicas
(1972)

Variable	Prueba estadística	Nivel de significancia
Valor de la propiedad (vivienda)	t= -1.40*	0.08
Valor de la vivienda	t= 1.46	0.07
Tipo de construcción de la vivienda (por punteo)	t= 3.60	0.001
Número de posesiones modernas (no esenciales)	t= 2.83	0.003
Salario diario	t= 1.94	0.02
Ingresos subsidiarios	t= -0.10	N.S.
Total de tierra sembrada, 1971	t= 1.65	0.06
Total de tierras de milpa propias, 1972	t= 0.28	N.S.
Total de tierras de café propias, 1972	t= 2.86	0.002
Valor del total de tierra propia, 1972	t= 2.36	0.01
Valor de la cosecha de maíz, 1972	t= 1.67	0.05
Valor de la cosecha de café, 1972	t= 2.85	0.003
Valor de la producción agrícola 1972	t= 2.29	0.01
Total de salario en efectivo, 1972	t= 2.53	0.006

* El signo menos (-) indica que los Bajos /No-participantes tuvieron un mayor punteo en la variable medida en particular.

Conclusiones

La evaluación se propone proporcionar al nivel de decisión política y técnica la información pertinente para el desarrollo de políticas y programas. La forma ideal para alcanzar estos objetivos es valiéndose del uso de diseños experimentales e incluyendo varias comunidades. Cuando ello no es posible, el evaluador debe adaptarse a las realidades de la situación, usando las herramientas disponibles de obtener, bajo las circunstancias prevalentes, las respuestas más apropiadas.

En el caso de una sola comunidad, esto requiere medidas detalladas y de repetición durante cierto tiempo. También necesita la clasificación de individuos o grupos hogareños autoseleccionados y sus reacciones a los estímulos del proyecto, lo que requiere la medición del impacto bruto y un análisis de los individuos o grupos de individuos que contribuyen, tanto positiva como negativamente, a los indicadores de impacto. En evaluaciones de una sola comunidad, este proceso de análisis y síntesis puede ayudar a responder a las preguntas de los diseñadores del programa. A pesar de ello, este tipo de evaluación nunca puede proporcionar respuestas de causa-efecto a un nivel muy específico.