

# Evaluación de impacto de programas de acompañamiento

ASCENETH MARÍA SASTRE CIFUENTES  
YOHANNA MARÍA MONTENEGRO MEJÍA\*

Uno de los mayores retos en materia de evaluación es determinar si los programas educativos o sociales cumplen con el fin para el cual fueron diseñados. Con este propósito, las evaluaciones de impacto proveen metodologías rigurosas para determinar si los programas alcanzan los resultados esperados en sus beneficiarios mediante el uso de indicadores adecuados (Consejo Nacional de Evaluación de la Política de Desarrollo Social [Coneval], 2012).

Esencialmente, una evaluación de impacto consiste en generar evidencias sobre qué políticas sociales funcionan y cuáles no, lo cual se puede hacer dentro de un marco clásico de evaluación de impacto comparando indicadores de resultados para un grupo de beneficiarios y un grupo gemelo que no participó en el programa. También se pueden realizar evaluaciones de impacto para examinar alternativas de implementación dentro de un programa o analizar diferentes programas para comparar su desempeño.

Al respecto Gertler, Martínez, Premand, Rawlings y Vermeersch (2010) proponen iniciar precisando las preguntas de evaluación y

---

\* Docentes investigadoras responsables del proyecto “Evaluación del impacto de los programas de apoyo relacionados con la permanencia estudiantil en la Universidad Santo Tomás de la modalidad presencial de pregrado, Bogotá, 2013”, en el marco del Convenio 626 USTA-Ministerio de Educación Nacional.

plantean diversos métodos, entre ellos, el de inferencia causal y las estimaciones contrafactuales, métodos de selección aleatoria, diseños de regresión continua, de diferencias y de pareamiento.

Estos autores subrayan algunas recomendaciones para mitigar los riesgos inherentes a la implementación de una evaluación de impacto: diseñar las evaluaciones al principio del ciclo del proyecto como parte del diseño y de la preparación del programa para planificar y recoger los datos de línea de base antes del comienzo del programa; complementar los resultados finales con otras informaciones, particularmente una evaluación de proceso y datos de monitores que ofrezcan una clara descripción de la implementación del programa, de modo que se entiendan las razones del logro de los objetivos, o en caso contrario, de los defectos en la implementación del programa o de sus deficiencias de diseño, y mantener identificadores comunes para las diferentes fuentes de datos de tal manera que puedan ser fácilmente vinculados durante los análisis.

Es importante también reconocer que las evaluaciones de impacto son útiles para aprender acerca del funcionamiento de los programas y comprobar alternativas programáticas que permitan poner a prueba componentes innovadores. Así, es posible incorporar innovaciones adicionales como pruebas piloto, en el contexto de una evaluación más grande que permita generar información valiosa que oriente decisiones futuras.

Es muy importante resaltar que las evaluaciones deben considerarse como un componente más de las operaciones de un programa y se deben dotar de personal y el presupuesto adecuado, con los recursos técnicos y financieros necesarios. Cuando la evaluación de impacto acompaña el diseño y la implementación de programas desde el principio, que es lo ideal, debe considerarse el tiempo de exposición que necesita la intervención para afectar los resultados con referencia a la línea de base, lo que según el programa podría tomar entre uno y cinco años. Por lo anterior, en su conjunto el ciclo completo de una evaluación de impacto acarrea al menos tres o cuatro años de trabajo y dedicación intensos, y exige contar con recursos financieros y técnicos adecuados en cada fase del proceso (Gertler *et al.*, 2010).

La evaluación de impacto ofrece respuestas concretas a preguntas específicas sobre políticas, que aportan información a las necesidades del organismo sobre el que estas se efectúan. También suministra información valiosa para actores de otras organizaciones que pueden aprender y tomar decisiones basadas en la evidencia generada por la evaluación. Así, una evaluación de impacto se convierte en parte del bien público global (Gertler *et al.*, 2010).

Como hasta acá se ha expuesto, la evaluación de impacto suele obedecer a un modelo de investigación de corte experimental que implica una toma de datos previa a la intervención que se compara con los datos generados posteriormente, y que puede enriquecerse con la comparación entre grupos gemelos en un diseño que garantiza condiciones de control.

Cuando se trata de evaluar el impacto de programas diversos que han venido ejecutándose desde diferentes momentos, con diversidad de historias y propuestas, como ha sido el caso de las experiencias de acompañamiento adelantadas en diferentes unidades académicas de la Universidad Santo Tomás, no son viables los estudios que suponen contar con líneas de base contrastables con datos posteriores a la intervención. En este sentido, las propuestas de evaluación de programas educativos, de corte más cualitativo, se convierten en la mejor alternativa para valorar los resultados de los programas, sin reñir con algunas de las recomendaciones derivadas de los estudios de impacto de programas sociales.

La propuesta de Stake (1975), citado por Stufflebeam y Shinkfield (1987), construida sobre la noción de Ralph Tyler acerca de que los evaluadores en educación deben comparar los resultados deseados y los observados en la línea de los estudios de impacto previamente descritos, amplía el concepto de *evaluación* pidiendo un examen de los antecedentes, del proceso, de las normas y de los juicios, además de los resultados.

Este modelo, conocido como evaluación respondiente, asume que las intenciones pueden cambiar y pide una comunicación continua entre el evaluador y los actores del programa, con el fin de descubrir, investigar y solucionar los problemas, haciendo de este un método pluralista, flexible, interactivo, holístico, subjetivo y orientado hacia el servicio.

En efecto, alejándose de los experimentos comparativos como método de evaluación, exige descripciones completas de los programas y subraya la importancia de la información subjetiva. En este modelo, las evaluaciones deben ayudar a los actores de los programas a observar y mejorar lo que están haciendo. Los evaluadores deben describir los programas con relación tanto a los antecedentes y a las operaciones como a los resultados, de modo que los efectos secundarios y los logros accidentales se incorporen en el estudio como los resultados buscados.

En lugar de presentar conclusiones finales resumidas, los evaluadores recopilan, procesan, analizan y reflejan los juicios de una amplia gama de personas interesadas en el objeto de la evaluación. En este sentido, los experimentos y tests regularizados resultan, cuando no inadecuados, insuficientes frente a la complejidad del proceso.

Stake (1975) ofrece un esquema para la recopilación de datos que explicita los tipos de información que el evaluador deberá recoger, analizar y presentar: se trata de información acerca de antecedentes, transacciones y resultados basándose en diversas fuentes, para así aproximarse al objetivo de abordar la totalidad de la evaluación (tabla 24).

Tabla 24. Esquema de Stake para la recopilación de datos en la evaluación de un programa educativo

	Intenciones (lo que pretende la gente)	Observaciones (lo que percibe la gente)		Normas (lo que la gente generalmente aprueba)	Juicios (lo que la gente cree que debe ser el programa principal)
Base lógica			Antecedentes		
			Transacciones		
			Resultados		
	Matriz de descripción			Matriz de juicio	

Fuente: elaboración propia.

Los antecedentes se refieren a la información más relevante sobre el historial, es decir, sobre las condiciones existentes antes del proceso formativo (ejecución del programa) y quizá relacionada con los resultados. Las transacciones incluyen los innumerables encuentros entre los estudiantes (beneficiarios del programa) con otras personas, profesores, padres de familia, administrativos, directivos y otros estudiantes en los que se disciernen las acciones del programa. Los resultados se refieren a lo que se consigue a través del programa: capacidades, logros, aptitudes, aspiraciones, el impacto que causa en todos los participantes, profesores, padres, administradores, estudiantes y otros; “Incluye los resultados evidentes y los confusos, los buscados y los no buscados, los de largo y los de corto alcance” (Stufflebeam y Shinkfield, 1987, p. 244).

Antecedentes, transacciones y resultados son los conceptos centrales que orientan lo que debe hacerse al describir y juzgar un programa, y son la dimensión vertical del esquema. La dimensión horizontal son las intenciones y las observaciones, para la matriz de descripción, y las normas y los juicios para la matriz de juicio.

Las intenciones se refieren a todo lo que se planea, incluso los antecedentes, las actividades de enseñanza y aprendizaje y los resultados deseados. Tanto lo que se incluye como lo que se excluye y que exprese los propósitos de los educadores en sus palabras. Las observaciones se refieren a los antecedentes, las transacciones y los resultados que han sido observados y anotados, que pueden haber sido recopilados a través de distintas fuentes e instrumentos de recolección de datos. Seguir esta lógica permite al evaluador reflexionar ampliamente, más allá de la confrontación de objetivos con resultados.

Stake (1975) recomienda investigar además la base lógica del programa, es decir, las intenciones filosóficas de este, y sus elementos conceptuales que, entre otras cosas, orientan la identificación de grupos de referencia que permitirían identificar las normas y aprobar los juicios.

Esta información descriptiva se somete a análisis de congruencia y a análisis de contingencia. El primero averigua si los propósitos se han cumplido, si son congruentes los antecedentes observados con los pronosticados verticalmente, mientras que el segundo investiga las contingencias entre las intenciones (contingencia lógica) y las

observaciones (contingencia empírica) horizontalmente. Estos últimos análisis de contingencia de las condiciones observadas deben estar basados en criterios de evidencia empírica, y requieren datos procedentes del programa que se está investigando.

La recopilación y el análisis de la información descriptiva y de la descripción de la base lógica del programa proporciona el fundamento de la matriz de juicios: la identificación de normas y la formulación de juicios acerca del mérito del programa. Las normas son los criterios explícitos para valorar la excelencia de una oferta educativa, mientras que los juicios son una decisión sobre la atención que hay que prestar a las normas de cada grupo de referencia a la hora de decidir si hay que emprender alguna acción administrativa. Hay normas absolutas (convicciones personales acerca de lo que es bueno y deseable para un programa) y normas relativas (características de los programas alternativos que se creen satisfactorios, y sugieren normas de calidad).

Así, el procedimiento para la generación de juicios implica, en primer lugar, la recopilación y el análisis de la información descriptiva y describir la base lógica del programa. Luego, la identificación de normas absolutas (convicciones formales e informales conducidas por importantes grupos de referencia acerca de qué normas de calidad deben utilizarse) y la recopilación de datos descriptivos de otros programas para extraer las normas relativas para compararlas con el programa en cuestión.

Con este material, el evaluador valora hasta qué punto el programa en cuestión cumple las normas absolutas y relativas y, finalmente, juzga el programa decidiendo qué normas toma en consideración, asignando a estas más o menos importancia. El producto final del proceso evaluativo, en todo caso, deberá conducir a la toma de decisiones y a desarrollar un conocimiento básico sobre la educación, no a la búsqueda de problemas.

En educación, afirman Stufflebeam y Shinkfield (1987), el punto en que empieza a actuar el evaluador está situado, por lo general, después de que el programa ya ha empezado, tal como ocurre en el caso que nos ocupa en la Universidad Santo Tomás con los programas de acompañamiento estudiantil, y su principal papel es proporcionar información evaluativa útil a quienes están operando con el programa,

y la evaluación respondiente de Stake (1975) subraya marcos para el desarrollo del aprendizaje, las transacciones didácticas, los datos para los juicios, el informe holístico y la asistencia a los educadores.

Para Stake (1975) la evaluación es un valor observado comparado con alguna norma, definición representada en la siguiente proporción:

Toda la constelación de valores de un programa  
Evaluación = Complejo de expectativas y criterios  
que distintas personas tienen sobre el programa

En suma, la evaluación respondiente de Stake (1975) se orienta a describir las actividades de un programa educativo, y concede más importancia a los problemas que a las teorías, toma en cuenta las diferentes interpretaciones de aquellos que están implicados en el programa, debe responder a las necesidades de información y al nivel de conocimiento de quienes se encuentran interesados en el programa y proporciona a los interesados la experiencia vicaria del programa y los implica en sus análisis y valoración.

El modelo de Stake (1975) ofrece una metodología que refleja la totalidad, la complejidad y la importancia de los programas y que corresponde en la investigación a un enfoque cualitativo de investigación participativa, compleja. Es, por tanto, un modelo apropiado para la evaluación del impacto de programas de acompañamiento estudiantil de la Universidad Santo Tomás.

