

Uso de evaluaciones homogéneas para evaluar programas sociales: Caso México

Hortensia Pérez Seldner¹

Resumen

MEDIR Y ANALIZAR LOS RESULTADOS OBTENIDOS DEBE SER EL OBJETIVO PRINCIPAL DE UN SISTEMA DE MONITOREO Y EVALUACIÓN. SIN EMBARGO, PARA LOGRAR LA MEDICIÓN DE LOS RESULTADOS FINALES O EL IMPACTO DE UNA INTERVENCIÓN ES NECESARIO CONTAR CON INFORMACIÓN REFERENTE A LAS ACTIVIDADES, LOS INDICADORES, LOS PROCESOS Y LOS RESULTADOS INTERMEDIOS DE LOS PROGRAMAS. SIN DUDA, CADA PROGRAMA SOCIAL ES DIFERENTE EN CUANTO A SU OPERACIÓN, METAS Y CONTEXTO; PERO EXISTEN ASPECTOS QUE PUEDEN SER ANALIZADOS DE FORMA HOMOGÉNEA ENTRE PROGRAMAS PARA REALIZAR VALORACIONES PRÁCTICAS Y MENOS COSTOSAS.

EN ESTE SENTIDO, EN MÉXICO SE HAN DISEÑADO DIVERSOS INSTRUMENTOS PARA ALCANZAR UNO O VARIOS DE LOS PRINCIPALES OBJETIVOS DE UNA EVALUACIÓN: MEJORAR EL PROGRAMA, BRINDAR INFORMACIÓN PARA LA TOMA DE DECISIONES PRESUPUESTARIAS; PROMOVER LA TRANSPARENCIA Y LA RENDICIÓN DE CUENTAS; REALIZAR UN ANÁLISIS INTEGRAL DEL SECTOR, ENTRE OTROS. DICHS INSTRUMENTOS INCLUYEN EVALUACIONES HOMOGÉNEAS QUE SE REALIZAN A LA MAYORÍA DE LOS PROGRAMAS SOCIALES.

POR LO TANTO, EL PRESENTE DOCUMENTO PRETENDE HACER UN ANÁLISIS SOBRE LAS FORTALEZAS Y DEBILIDADES DE LAS EVALUACIONES HOMOGÉNEAS CON BASE EN LOS RESULTADOS OBTENIDOS EN MÉXICO, Y FOMENTAR ASÍ LA DISCUSIÓN SOBRE LA UTILIDAD DE ESTE TIPO DE PRÁCTICAS Y LOS POSIBLES RETOS A SUPERAR.

Abstract

MEASURING AND ANALYZING ULTIMATE OUTCOMES MUST BE THE MAIN OBJECTIVE OF A MONITORING AND EVALUATION SYSTEM. HOWEVER, IN ORDER TO GET THERE IT IS NECESSARY TO COLLECT DATA ABOUT PROGRAMS' ACTIVITIES, INDICATORS, PROCESSES AND INTERMEDIATE OUTCOMES. UNDOUBTEDLY, EACH SOCIAL PROGRAM IS DIFFERENT IN TERMS OF PROCESSES, GOALS AND CONTEXT; BUT THERE ARE FACTORS THAT COULD BE HOMOGENEOUSLY ANALYZED AMONG PROGRAMS SO AS TO DEVELOP PRACTICAL AND LESS COSTLY ASSESSMENTS.

IN THIS REGARD, SEVERAL INSTRUMENTS HAVE BEEN DESIGNED IN MEXICO TO ACHIEVE ONE OR MORE OF THE MAIN OBJECTIVES OF AN EVALUATION: IMPROVE PROGRAMS; PROMOTE INFORMED DECISION-MAKING; PROMOTE TRANSPARENCY AND ACCOUNTABILITY; DEVELOP AN INTEGRAL ANALYSIS OF THE FIELD, ETC. THESE INSTRUMENTS INCLUDE HOMOGENEOUS EVALUATIONS APPLIED TO THE MAJORITY OF SOCIAL PROGRAMS.

THUS, THIS PAPER AIMS TO ANALYZE THE STRENGTHS AND WEAKNESSES OF HOMOGENEOUS EVALUATIONS, BASED ON THE RESULTS OF THE MEXICAN CASE, AND ENCOURAGE THE DISCUSSION ABOUT THE CONVENIENCE OF THESE PRACTICES AND THE POTENTIAL CHALLENGES.

PALAVRAS-CHAVE:

Evaluación; Programas sociales; Evaluación homogénea

Introducción

Medir y analizar los resultados obtenidos debe ser el objetivo principal de un Sistema de Monitoreo y Evaluación. Sin embargo, para lograr la medición de los resultados finales o el impacto de una intervención es necesario contar con información referente a las actividades, los indicadores, los procesos y los resultados intermedios de los programas.

Sin duda, cada programa social es diferente en cuanto a su operación, metas y contexto, pero existen aspectos que pueden ser analizados de forma homogénea entre programas para realizar valoraciones prácticas y menos costosas.

En México, el gobierno federal opera alrededor de 270 programas y acciones de desarrollo social², que cada año necesitan ser analizados para tomar decisiones informadas. En la práctica, las instituciones no tienen los recursos (financieros, humanos, materiales) necesarios para desarrollar evaluaciones de impacto a la totalidad de sus programas, pero eso no debe desvigorizar los esfuerzos por conocer el desempeño y productividad de cada una de las intervenciones que funcionan con recursos públicos.

En México, el Sistema de Monitoreo y Evaluación (ver Anexo 1) contempla ocho principales tipos de evaluación, que consideran diferentes actores, usuarios, objetivos y fre-

cuencias de realización. Los tipos de evaluación son los siguientes³:

1) *Evaluación de Diseño (ED)*: Se enfoca en la consistencia y lógica interna de los programas. Se realiza mediante gabinete con la información proporcionada por los programas. De acuerdo con el Programa Anual de Evaluación (PAE) 2013 estas evaluaciones deben realizarse a todo programa que se encuentre en su primer año de operación o que haya tenido cambios sustanciales. Los principales usuarios son los programas mismos y los tomadores de decisiones, en especial los que impulsaron la nueva intervención.

2) *Evaluación de Consistencia y Resultados (ECR)*: Su principal objetivo es realizar un diagnóstico sobre la capacidad institucional y de gestión de los programas. El ejercicio es de gabinete pero con entrevistas y constante comunicación con los operadores del programa; sus principales usuarios son los programas y los tomadores de decisiones. En México esta evaluación se ha desarrollado en dos ocasiones de manera simultánea a la mayoría de los programas sociales.

3) *Evaluación de Procesos (EP)*: Realiza un análisis de los procesos operativos del programa y su contribución a los objetivos principales del mismo. Esta

evaluación podría servir como preámbulo para una evaluación de impacto, ya que permite identificar los procesos, los sistemas y la información disponible. El principal usuario es el programa.

4) *Evaluación Específica de Desempeño (EED)*: Es una valoración sintética de la información disponible del programa. Se realiza periódicamente de acuerdo con lo establecido en los programas anuales de evaluación⁴. Los principales usuarios de este análisis son los tomadores de decisiones (Secretarios, directivos de los programas, la Secretaría de Hacienda y Crédito Público, la Cámara de Diputados, entre otros).

5) *Evaluación de Indicadores (EI)*: Analiza la pertinencia y alcance de los indicadores de un programa.

6) *Evaluación Estratégica*: Realiza una valoración de las políticas y estrategias de desarrollo social. Generalmente abarcan

más de un programa o intervención. Entre las evaluaciones estratégicas que ha coordinado el CONEVAL se encuentran las siguientes: Protección Social en México, Ramo 33, Nutrición y Abasto, Microcréditos, Mortalidad Materna, entre otros temas.

7) *Evaluación Complementaria*: Este tipo de evaluaciones son de aplicación opcional según las necesidades e intereses de las dependencias y entidades, con la finalidad de mejorar su gestión y obtener evidencia adicional sobre su desempeño.

8) *Evaluación de Impacto*: El objetivo es medir los efectos netos del programa. Al identificar efectos atribuibles a la intervención, esta evaluación es de utilidad tanto para el programa como para los tomadores de decisiones y el público en general.

Para definir el tipo de evaluación que necesita un programa de desarrollo social se debe considerar tanto su nivel de madurez como el

2 De acuerdo con las definiciones del Inventario CONEVAL de Programas y Acciones Federales de Desarrollo Social 2011-2012, en México existen 179 programas y 94 acciones.

3 Las definiciones son elaboración propia con base en la información contenida en la página web del CONEVAL: www.coneval.gob.mx y en los Lineamientos Generales para la Evaluación de Programas Federales.

4 A la fecha la EED se ha llevado a cabo en 131 programas en 2008-2009, 127 programas en 2009-2010, 133 programas en 2010-2011 y está por concluir la EED 2012-2013 para 137 programas.

uso específico que se le dará a la información y resultados obtenidos: mejorar el programa, brindar información para la toma de decisiones presupuestarias, transparencia, rendición de cuentas, análisis integral del sector, entre otros. Asimismo, es necesario contar con los sistemas de información pertinentes para facilitar las mediciones requeridas.

En este sentido, el Consejo Nacional de Evaluación de la Política de Desarrollo Social (CONEVAL)⁵ ha diseñado diversos instrumentos para alcanzar uno o varios de los objetivos planteados, los cuales incluyen evaluaciones homogéneas que se realizan en conjunto a la mayoría de los programas con términos de referencia definidos.

Por lo tanto, el presente documento pretende hacer un análisis sobre las fortalezas y debilidades de las evaluaciones homogéneas con base en los resultados obtenidos en México, y fomentar así la discusión sobre la utilidad de este tipo de prácticas y los posibles retos a superar.

Evaluaciones homogéneas en México

A partir de su creación en 2004 el CONEVAL ha desarrollado diversos instrumentos de evaluación con términos de referencia homogéneos para programas sociales. Entre ellos destacan la Evaluación de Diseño, la Evaluación de Consistencia y Resultados, la Evaluación Específica de Desempeño y la Evaluación de Procesos.

Un ejemplo destacado de este tipo de ejercicios es la *Evaluación de Consistencia y Resultados (ECR)*, que evalúa sistemáticamente el diseño y desempeño de los programas federales, ofreciendo un diagnóstico sobre su capacidad institucional, organizacional y de gestión, con una orientación hacia resultados. La finalidad de esta evaluación es proveer información que retroalimente el diseño, la gestión y los resultados de los programas.

La ECR se implementó por primera vez a un total de 106 programas al principio del sexenio anterior, durante 2007-2008, y después de analizar los resultados y recibir retroalimentación por parte de los involucrados se realizó una revisión exhaustiva del instrumento con el principal objetivo de homogenizar la calidad de las evaluaciones. Dicha revisión comprendió grupos focales, análisis de construcción de preguntas y respuestas y evaluaciones piloto a lo largo de tres años, lo que derivó en un nuevo modelo de términos de referencia. El nuevo instrumento se aplicó al final del sexenio a un total de 131 programas sociales.

La ECR consta de seis temas a evaluar: (1) Diseño, (2) Planeación y Orientación a Resultados, (3) Cobertura y Focalización, (4) Operación, (5) Percepción de la Población Atendida y (6) Medición de Resultados.

Los términos de referencia están detalladamente definidos y cuentan con 51 preguntas

en total, de las cuales 34 son binarias (SI/NO) y si la respuesta es afirmativa, entonces se debe elegir un nivel de respuesta de acuerdo con criterios acumulativos establecidos para cada caso y fundamentar debidamente la respuesta seleccionada. De igual forma, en cada pregunta (binaria o no) se especifican elementos mínimos a considerar para dar respuesta, así como las fuentes mínimas de información que se deben revisar y, asimismo, se indican las preguntas relacionadas para apoyar al equipo evaluador en la preservación de la congruencia de sus respuestas. En el Anexo 2 se presenta un ejemplo de una pregunta binaria con los componentes antes descritos.

A su vez, la ECR solicita 20 anexos específicos que permiten la recopilación de información homogénea para análisis integrales adicionales, tales como la evolución de la cobertura, el análisis de coincidencias y complementariedades entre programas, el seguimiento a recomendaciones, entre otros.

Un aspecto importante para la operación de evaluaciones homogéneas es definir detalladamente el proceso de elaboración y revisión, así como si se utilizará un sistema de información también homogéneo. Para el caso de las ECR, el CONEVAL diseñó el Módulo para las Evaluaciones de Consistencia y Resultados (MOCYR), donde se desarrollaron la totalidad de las evaluaciones (contratadas o no directamente por la institución).

Al finalizar el ejercicio de la ECR 2011-2012 se realizó un estudio para conocer la percepción de los usuarios con respecto de la evaluación, representados por unidades de evaluación de las dependencias o entidades con programas evaluados, responsables de la operación de los programas y evaluadores externos que realizaron ECR.

El estudio analiza seis variables latentes: (1) utilidad del instrumento, (2) formato de la presentación, (3) mecánica, (4) evaluador externo, (5) contenido y (6) quejas.

5 De acuerdo con lo establecido en el Artículo 81 de la Ley General de Desarrollo Social el CONEVAL es un organismo es un organismo público descentralizado, con autonomía y capacidad técnica para generar información objetiva sobre la situación de la política social y la medición de la pobreza en México. Entre sus principales funciones está la de normar y coordinar la evaluación de programas y políticas sociales. (LGDS, 2004)

De acuerdo con Vargas y Merino (2012) el nivel de satisfacción de este tipo de evaluaciones ha ido aumentando en cada ejercicio y actualmente es un nivel relativamente alto para los estándares del sector público. Sin embargo, es interesante que existen percepciones recurrentes tales como que la evaluación es rígida en formato y contenido; que a pesar de cierta mejora, aún se requiere reforzar la comunicación entre el equipo evaluador y el evaluado; que los programas solicitan mayor retroalimentación en los resultados; que hay confusiones sobre que documentos pueden ser considerados para el desarrollo de la evaluación; también retoman la constante inquietud de que se deben impulsar otras evaluaciones como las de impacto, entre otros aspectos que permitirán mejorar tanto el instrumento como el sistema de evaluación en general.

Considerando la evolución de la ECR como instrumento homogéneo se pueden identificar las siguientes fortalezas con sus respectivos retos:

a. Permite la realización de un diagnóstico completo, que incluye los principales aspectos que todo programa debe contemplar. A pesar de buscar fines distintos, toda intervención debe contar con un diseño sólido y una operación orientada a resultados. Si bien es cierto que las preguntas binarias son rígidas en su definición también es importante señalar que dicha rigidez permite homologar los criterios cuando se están comparando justamente programas tan diversos.

El reto es encontrar el balance entre criterios detalladamente definidos y flexibilidad en las respuestas. Entre mejor se comprenda el instrumento por todas las partes involucradas (programa, coordinadores, evaluadores, usuarios) será más fácil lograr valoraciones objetivas que permitan comparaciones entre programas y que, a su vez, realcen los aspectos específicos de cada uno.

b. Permite conocer la institucionalización de las operaciones de cada programa. Es decir, identifica si el programa cuenta con la documentación requerida para seguir funcionando en caso de que cambien los responsables de operarlo. Asimismo, apoya en la organización de la información institucional existente.

Identificar la información oficial sigue siendo un reto, ya que el nivel de institucionalización varía entre dependencias o entidades, e incluso entre programas dentro de una misma dependencia. El objetivo debe ser impulsar la documentación no por burocracia sino por acoger las buenas prácticas y hacerlas parte de las operaciones cotidianas.

c. Al conocer los criterios para cada aspecto evaluado, el programa puede comprender con mayor claridad hacia dónde debe enfocar sus esfuerzos de mejora. Por ejemplo, si tiene definida y cuantificada su población objetivo, pero no tiene planeado un plazo para su revisión y actualización, entonces no alcanza todos los puntos posibles, pero

identifica la necesidad de reforzar su planeación para ejercicios futuros.

Es importante motivar al equipo evaluador para que fundamente a detalle (y de manera sintética) qué criterios faltan por lograr y que proponga acciones que permitan alcanzar dichos aspectos.

d. Permite la comparabilidad entre programas similares y la retroalimentación por medio de buenas prácticas compartidas.

El sólo hecho de contar con el mismo tipo de información para varios programas sociales permite realizar análisis comparativos, tal vez descriptivos, pero que pueden apoyar en la identificación de similitudes y complementariedades entre programas. Lo anterior, con el fin de compartir información e intercambiar buenas prácticas e ideas de mejora.

Cabe mencionar que la ECR incluye un formato exclusivo para identificar programas similares o complementarios, que al ser realizado por un externo tal vez no sea del todo preciso pero que puede servir de base para identificar aquellas intervenciones que atienden a la misma población, realizan actividades similares o persiguen objetivos comunes. Esta información puede ser de gran utilidad para análisis integrales.

e. Para los tomadores de decisiones (directivos, secretarios, diputados) el instrumento permite contar con información homogénea y fácil de comprender. Los resultados obtenidos son utilizados durante el proceso presupuestario, tanto por la Secretaría de Hacienda y Crédito Público (SHCP) como por los programas en específico.

Un ejemplo específico del uso de evaluaciones homogéneas es el Modelo Sintético de Información del Desempeño (MSD) que publica la SHCP. Para la variable de Evaluación de los programas sociales el MSD utiliza como fuente la *Valoración General del Desempeño de los Programas de Desarrollo Social* que elabora el CONEVAL⁶. (SHCP, 2012)

Sin embargo, el reto es no concentrarse sólo en las preguntas binarias por ejemplo (que por practicidad son más rápidas de analizar) sino también en las preguntas abiertas que brindan información cualitativa relevante de cada programa. Las discusiones y análisis deben contemplar una valoración sintética pero integral de cada programa.

f. El hecho de que la evaluación solicite cierta información o formatos específicos es un incentivo para que los responsables de los programas fortalezcan sus sistemas de información e intenten re-

6 La Valoración del Desempeño de los Programas de Desarrollo Social se puede encontrar en el siguiente vínculo del CONEVAL: http://www.coneval.gob.mx/Evaluacion/Paginas/Valoraci%C3%B3n_deL_desempe%C3%B1o_de_los_programas_2011-2012.aspx

copilar aquella documentación relevante que es requerida en la evaluación, y que les es de utilidad para la operación y análisis de su programa.

El reto es continuar retroalimentando los formatos y las características de la información solicitada con la finalidad de que sea útil y factible de recopilar.

g. El costo es accesible y la capacitación para la aplicación del instrumento está a cargo del CONEVAL, quien funge como apoyo durante todo el proceso.

Comparado con evaluaciones de impacto el costo es sustentable para la mayoría de los programas. De igual manera, es posible que conforme el sistema de evaluación y monitoreo siga madurando el mercado de evaluadores externos continúe ampliándose y los precios se ajusten.

Como se mencionó anteriormente, la capacitación para la aplicación, coordinación y elaboración de la ECR y cualquier tipo de evaluación homogénea es esencial para el óptimo desarrollo y aprovechamiento de este tipo de instrumentos.

Uno de los principales retos es la aceptación por parte de los responsables de

los programas, ya que al ser ellos los que mejor conocen sus programas les resulta difícil ser catalogados como uno más del conjunto de intervenciones sociales. Sin embargo, este paradigma puede ser modificado con sesiones de capacitación y sensibilización que les permitan comprender las ventajas del instrumento y reconocer que el programa mismo es el principal usuario.

Así como la ECR, en México también existen términos de referencia establecidos para las Evaluaciones de Diseño y para las Evaluaciones Específicas de Desempeño, que en conjunto han logrado la participación de todos los actores involucrados, la mejora de programas, el aumento de la transparencia y rendición de cuentas, así como el gradual incremento en el uso de información técnica en la toma de decisiones.

El reto en general es continuar fortaleciendo el Sistema de Evaluación y Monitoreo mexicano, diagnosticar las necesidades de evaluación de cada programa y seguir impulsando el desarrollo de evaluaciones homogéneas para resultados intermedios en conjunto con la planeación de otro tipo de evaluaciones.

■ ANEXO 1: DISEÑO DEL SISTEMA DE MONITOREO Y EVALUACIÓN EN MÉXICO



CONEVAL (2011). Resultados 2006-2009 del CONEVAL

■ ANEXO 2: EJEMPLO DE PREGUNTA BINARIA EN LA EVALUACIÓN DE CONSISTENCIA Y RESULTADOS (ECR)

1. Existe un diagnóstico del problema que atiende el programa que describa de manera específica:
 - a) Causas, efectos y características del problema.
 - b) Cuantificación, características y ubicación territorial de la población que presenta el problema.
 - c) El plazo para su revisión y su actualización.

Si el programa no cuenta con un diagnóstico del problema al que atiende, se considera información inexistente y, por lo tanto, la respuesta es "No".

Si cuenta con información para responder la pregunta, es decir, si la respuesta es "Sí" se debe seleccionar un nivel según los siguientes criterios:

En la respuesta se deben incluir las principales causas y los efectos del problema señalados en el diagnóstico. Adicionalmente, se debe valorar la vigencia del diagnóstico y, en su caso, se pondrán sugerencias para mejorarlo.

Las fuentes de información mínimas a utilizar deben ser documentos de diagnóstico y árbol de problema.

La respuesta a esta pregunta debe ser consistente con las respuestas de las preguntas 1, 3 y 7.

NIVEL	CRITERIOS
1	El programa cuenta con un diagnóstico del problema, y El diagnóstico no cuenta con las características establecidas en la pregunta.
2	El programa cuenta con un diagnóstico del problema, y El diagnóstico cuenta con una de las características establecidas en la pregunta.
3	El programa cuenta con un diagnóstico del problema, y El diagnóstico cuenta con dos de las características establecidas en la pregunta.
4	El programa cuenta con un diagnóstico del problema, y El diagnóstico cuenta con todas las características establecidas en la pregunta.

Referencias

- CASTRO, M.; López-Acevedo, G.; Beeker, Busjeet G.; Fernández Ordonez, X. (2009). El Sistema de M&E en México: Un salto del nivel sectorial al nacional. Serie de Documentos de Trabajo para el Desarrollo de la Capacidad de Evaluación. Banco Mundial. No. 20.
- CONEVAL (2011). Resultados 2006-2009 del CONEVAL.
- CONEVAL. (2012). Modelo de Términos de Referencia de la Evaluación de Consistencia y Resultados.
- CONEVAL. (2013). Evaluaciones y resultados de programas. Consultado el 29 de julio de 2013 en: <http://www.coneval.gob.mx/Evaluacion/Paginas/Evaluaciones-y-resultados-de-programas.aspx>
- CONEVAL. Valoración del Desempeño de los Programas 2011-2012. Evaluación y Monitoreo de Programas Sociales. Consultado el 05 de agosto de 2013 en:
http://www.coneval.gob.mx/Evaluacion/Paginas/Valoraci%C3%B3n_del_desempe%C3%B1o_de_los_programas_2011-2012.aspx
- LEY General de Desarrollo Social. (2004). Texto vigente de la nueva ley publicada en el Diario Oficial de la Federación el 20 de enero de 2004.
- LINEAMIENTOS Generales para la Evaluación de Programas Federales. (2007). Texto vigente publicado en el Diario Oficial de la Federación el 30 de marzo de 2007.
- SECRETARÍA DE HACIENDA Y CRÉDITO PÚBLICO SECRETARÍA DE LA FUNCIÓN PÚBLICA CONSEJO NACIONAL DE EVALUACIÓN DE LA POLÍTICA DE DESARROLLO SOCIAL. (2013) Programa Anual de Evaluación para el Ejercicio Fiscal 2013 de los Programas Federales de la Administración Pública Federal.
- SECRETARÍA DE HACIENDA Y CRÉDITO PÚBLICO. (2012) Modelo Sintético de Información del Desempeño (MSD). Anexo Metodológico. Unidad de Evaluación del Desempeño. Consultado el 30 de julio de 2013 en:
<http://www.transparenciapresupuestaria.gob.mx/ptp/ServletImagen?tipo=pdf&idDoc=242>
- Vargas, D.; Merino, M. (2012). Medición de Satisfacción de Usuarios de las Evaluaciones de Consistencia y Resultados 2011-2012. Documento elaborado para el CONEVAL.