



Anexo IV. Los desafíos y dilemas asociados a la evaluación de la efectividad de instrumentos de políticas públicas

Este documento ha sido elaborado en cooperación con la Comisión Nacional para el Uso Eficiente de la Energía (Conuee).

Apoiado por

European Union Energy Initiative

Partnership Dialogue Facility (EUEI PDF)



Y el

Programa de Energía Sustentable en México

Implementado por Encargo del Ministerio Federal de Cooperación Económica y Desarrollo (BMZ)



c/o Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH

P.O. Box 5180, 65726 Eschborn, Germany

info@euei-pdf.org

www.euei-pdf.org

Autores

Daniel Bouille, Claudio Carpio, Nicolás Di Sbroiavacca, Hilda Dubrovsky, Gustavo Nadal, Francisco Lallana, Raul Landaveri, Héctor Pistonesi, Jorge Plauchú, Marina Recalde, y Rafael Soria (Fundación Bariloche).

Con comentarios y contribuciones de:

Odón de Buen, Juan Ignacio Navarrete, Israel Jáuregui, Pedro Hernández, Flor Chávez, Oscar Ruiz e Ilse Ávalos (Conuee).

Ernesto Feilbogen, Daniela Méndez, Fairuz Loutfi, Veronica Gómez (GIZ México).

Coordinación editorial: Ángel Azamar, Daniela Méndez y Veronica Gómez (GIZ México).

Ciudad de México, septiembre 2018

El Motor de Diálogo y Cooperación (EUEI PDF) es un instrumento de la EU Energy Initiative (EU EI). EUEI PDF actualmente recibe contribuciones de la Comisión Europea, Alemania, Austria, Finlandia, Italia, los Países Bajos y Suecia.



1. Los desafíos y dilemas asociados a la evaluación de la efectividad de instrumentos de políticas públicas

Para poder examinar y evaluar las reacciones/respuestas de los industriales a los instrumentos de política de eficiencia energética es necesario conocer la racionalidad del sector, es decir, las características de su comportamiento en la toma de decisiones.

Es claro que esto ha de diferir con el sector de actividad, el tamaño de la empresa (pequeña, mediana, grande) y otros múltiples factores de contexto. Para poder investigarlo es necesario realizar encuestas representativas por medio de un muestreo aleatorio (realizar un censo sería muy costoso y llevaría mucho más tiempo).

Ya sea que se opte por un medio u otro lo cierto es que habremos podido conocer ese comportamiento (siempre con cierto margen de error) en un momento del pasado más o menos reciente, en términos de ciertos valores medios y con determinados valores de dispersión.

Es decir que el resultado obtenido no sería de fácil aplicación en los individuos (empresas) que componen el universo o los sub-universos (en el sentido estadístico de esas nociones). Esto último debido a tres razones fundamentales:

- En primer lugar, porque el comportamiento de dicha unidad empresaria puede diferir considerablemente del comportamiento promedio que se ha estimado.
- En segundo término, porque el comportamiento futuro de ese actor no habrá de ser necesariamente (lo más probable es que no lo sea) similar al registrado en el pasado (en el tiempo que se realizó la investigación previamente descrita), máxime que lo que se pretende es analizar el efecto de una intervención de política que no existía o que es diferente de la situación previamente existente. No sería posible extrapolar al futuro comportamientos del pasado.
- Por último, la estadística no puede establecer relaciones de causalidad esencial o filosófica. Es decir, existen nociones de causalidad estadística entre variables, aplicables a un conjunto específico de datos, pero dicha causalidad no es aplicable necesariamente a otro conjunto de datos. Es decir que dicha causalidad no es de aplicación universal.

Una aproximación diferente al tema es plantear el problema como un experimento. Si se tratara de un experimento controlado de un fenómeno serían necesarios los siguientes datos:

- a) La especificación de todos los factores que pueden tener un efecto mensurable sobre el fenómeno a investigar debe ser conocida a priori. Se trata de las condiciones iniciales o las proposiciones explicativas de todos los elementos con influencia medible sobre el fenómeno bajo estudio o investigación.
- b) La proposición que se pretende probar del tipo “si A entonces B” en el caso del instrumento indirecto de política de eficiencia energética sería, por ejemplo: “Si se aplica un subsidio de tal magnitud entonces la empresa habrá de responder con tal grado de ahorro energético”.
- c) La descripción del fenómeno empírico a ser explicado que debe ser verificable.

Este tipo de experimentos constituyen sistemas cerrados donde todo se encuentra bajo el control del investigador y representa una imagen ideal de laboratorio.

En el caso de las ciencias físico-naturales los experimentos sólo se aproximan a estas condiciones ideales. A este respecto hay que recordar el Principio de Indeterminación de Heisenberg: *en el intento de medir puede alterarse el fenómeno a medir, con lo que se produce una incertidumbre acerca de lo que se está midiendo.*¹

En el ámbito de las ciencias sociales, como es el caso de la economía, no es posible realizar este tipo de experimentos, donde se puedan controlar las condiciones o factores que pueden tener un efecto medible sobre el fenómeno a estudiar. Por tanto, las proposiciones que se plantean son del tipo siguiente:

“Si A, *ceteris paribus* entonces B”. La cláusula “*ceteris paribus*” indica que todos los factores no considerados no se modifican.

En la medida en que no se especifiquen todos aquellos factores que engloba la cláusula “*ceteris paribus*” la anterior proposición no podrá ser falsificada o confirmada empíricamente ya que todos los resultados desfavorables pueden siempre atribuirse a modificaciones de aquella cláusula. Es decir, a un cambio en las condiciones de contexto.

Estas dificultades de verificación son en gran parte la causa de la proliferación de teorías y modelos sin demasiada preocupación por su relevancia empírica. De este modo, la economía aparece como una disciplina casi carente de "hechos". Sus teorías son rara vez cuestionadas aún si múltiples hechos vienen en su contra.

¹ El físico alemán **Werner Karl Heisenberg** desarrolló en 1927 el **principio de incertidumbre** ante la dificultad de expresar en lenguaje matemático la relación entre la posición del electrón y su momento lineal (velocidad masa), el cual explica que no es posible medir al mismo tiempo con exactitud donde se encuentra el objeto y cuán rápido se está moviendo.

De acuerdo con la filosofía de la ciencia, en el caso de una disciplina empírica una teoría no puede proporcionar una explicación científicamente válida si no es susceptible de ser corroborada empíricamente. Es decir, si no es posible deducir de dicha teoría proposiciones cuyo valor de verdad pueda establecerse por medio de un procedimiento adecuado de observación.

Tabla 1: Paradigmas y modelos vinculados a evaluación de políticas públicas

Egon Guba, en su libro “The Paradigm Dialog”², distingue cuatro grandes paradigmas en la indagación social:

Positivismo:

- a) La realidad existe y se rige por leyes de causa-efecto que pueden ser conocidas.
- b) La indagación puede estar libre de valores.
- c) Las hipótesis pueden ser comprobadas empíricamente.

Post-positivismo:

- a) La realidad existe, pero no puede ser enteramente comprendida o explicada. Hay una multiplicidad de causas y efectos.
- b) La objetividad es un ideal, pero requiere de una comunidad crítica.
- c) Es crítico con el experimentalismo, y enfatiza las aproximaciones cualitativas, la teoría y el descubrimiento.

Teoría crítica:

- a) La realidad existe, pero no puede ser enteramente comprendida o explicada. Hay una multiplicidad de causas y efectos.
- b) Los valores median en la indagación.
- c) Propone la eliminación de una falsa conciencia y facilita y participa en la transformación.

Constructivismo:

- a) La realidad existe como constructo mental relativo a aquellos que lo sostienen.
- b) El conocimiento y el conocedor son parte de una misma entidad subjetiva.
- c) Identifica, compara y describe las varias construcciones que existen (hermenéutico y dialéctico).

En cuanto a las “evaluaciones” pueden distinguirse cuatro “enfoques”, como puede verse en la siguiente tabla:

² Guba, Egon G. “The Paradigm dialog” – Sage Publications – Londres – 1990.

Tabla 2: Enfoques de evaluación de políticas públicas

Políticas públicas	
Experimental	Se basa en perspectiva positivista que consiste en aplicar la metodología de las ciencias naturales a las intervenciones públicas. En los planes experimentales se comparan grupos extraídos al azar que tienen características similares y a uno de ellos se le aplica la intervención y al otro no. El mayor empeño del diseño experimental será tratar de eliminar todas las variables que puedan afectar la validez de los resultados, ya que la atribución de los efectos observados como causa de la intervención se basará en la hipótesis de que no existen otras causas posibles imputables, que hayan afectado los resultados.
Pragmática	Pone el acento, sobre todo, en la utilidad y las constataciones de la evaluación que debe ser adaptada a los objetivos y a las prácticas de los principales decisores: la factibilidad política, la oportunidad y el costo son importantes. Se utiliza, generalmente, como herramienta de evaluación de los planes “cuasi-experimentales” -en los que se enfrentan diferentes explicaciones, posibles resultados, que se descartan por otros medio que la extracción al azar, de los estudios de casos y diversos métodos descriptivos.
Económica	Se sitúa, a menudo, aparte de la evaluación general. Aplica la idea denominada de racionalidad-económica. El análisis costo-beneficio y evaluación costo-eficacia o costo-efectividad, figuran entre las herramientas empleadas. Produce información sobre los costos del programa como criterio central de evaluación.
Naturista	Rechaza tanto la evaluación experimental como la económica, afirmando que la evaluación se construye socialmente y cambia por la acción de los individuos. En consecuencia, la evaluación no puede aportar respuestas exactas y objetivas, sino que es, a lo más, un agente que facilita el consenso entre las partes implicadas. Los métodos empleados, como la Etnografía, los estudios de Observación-participante, los Estudios de Casos, se dirigen a describir y clarificar los valores y presupuestos de las partes implicadas y prestan ayuda a la negociación del proceso de evaluación.

En el campo de la realidad social el método científico riguroso, detallado, universal y “estandarizable”, no existe. Cualquier método conlleva errores y puede encubrir conocimientos que no son tales. Más bien pueden identificarse dos paradigmas.

- El paradigma cuantitativo, que se caracteriza por la concepción global asentada en el positivismo lógico, el uso del método hipotético deductivo, su carácter particularista, orientado a los resultados. El supuesto de la objetividad se correspondería con las ciencias naturales.
- El paradigma cualitativo, asume una postura fenomenológica global, inductiva, estructuralista, subjetiva, orientada al proceso y propia de todas las disciplinas que tienen como tema de estudio la dimensión psicosocial de lo humano. Propia de la antropología social.

A cada uno de ellos se asocian métodos y técnicas. En el primer caso, métodos y técnicas cuantitativas como diseños experimentales y cuasiexperimentales; investigación por encuesta; cuestionarios estandarizados; registros estructurados de observación; técnicas estadísticas de análisis de datos, etc.

En el segundo caso, métodos y técnicas cualitativas: teoría fundamentada en la realidad; investigación histórica; investigación fenomenológica; historia de vida; investigación-acción.

Por lo anterior, se pueden identificar diversos métodos de evaluación de políticas públicas, como se muestra en la Tabla 3.

Tabla 3: Métodos de evaluación de políticas públicas

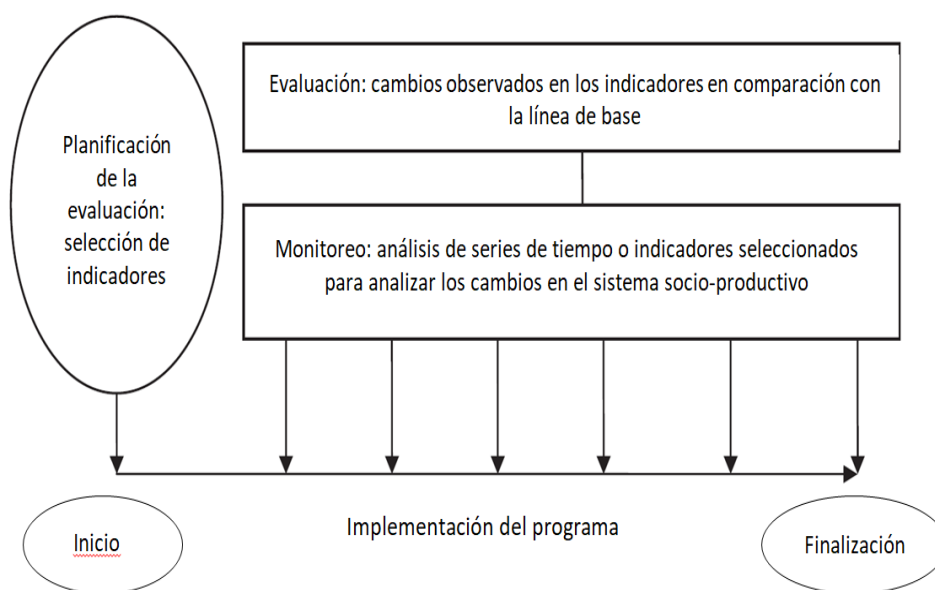
Evaluación de procesos	Evaluación de productos	Evaluación de resultados	Evaluación de impacto	Evaluación Costo-Beneficio	Evaluación institucional
Examina procedimientos, tareas y actividades	Analiza los bienes y servicios del programa o política	Analiza los cambios en los beneficiarios como resultado de los productos entregados	Analiza el impacto entre grupo de tratamiento y grupo de comparación	Mide tanto costos como beneficios en términos monetarios	Es una evaluación integral y holística sobre la institucionalidad que soporta el programa o política
Identifica insumos	Costo-eficiencia	Objetivos “Smart”	Método experimental	VAN y TIR	Fortalezas y debilidades
Identifica responsables	Examina calidad de los productos	Indicadores “crema”	Métodos cuasi-experimentales		Niveles de coordinación
Identifica productos Identifica beneficiarios Identifica cronograma	Verifica metas e indicadores para cada producto	Efectos directos sobre población objetivo	“Que habría pasado si el programa no hubiera existido”		¿La institución cumple con los objetivos?

Frente a políticas y estrategias que propongan un abanico de instrumentos, cada uno de ellos podría ser sometido a todos o varios de los métodos de evaluación para definir su aplicabilidad, de acuerdo con los objetivos que se persigan, las circunstancias que se enfrenten y las condiciones de entorno, habilitantes y barreras identificadas.

Finalmente, ha sido tal la influencia del positivismo en los procesos de evaluación, que en el campo de las ciencias sociales se han venido utilizando modelos que se han trasladado de proyectos de ciencias exactas, lo que ha generado una serie de resistencias de los científicos sociales, así como dificultades, puesto que esos modelos, son poco adaptables en los escenarios sociales donde los proyectos tienen lugar.³

En términos generales en la literatura predominan como propuestas de evaluación el enfoque experimental (con un grupo objetivo y uno de control)⁴, que obviamente en este caso, no sería muy útil; y el enfoque de impacto⁵ que padece seriamente del problema de causalidad. También se habla de evaluación cuando en realidad se trata de control o monitoreo (M&C) por medio de indicadores que es lo que parece más razonable⁶. En el caso de EE. UU., lo que predominó fue un enfoque muy influenciado por el positivismo y un optimismo metodológico en lo que se refiere a objetividad de la evaluación (que predominó en las décadas de los años 60 y 70), luego de cambios políticos⁷, la evaluación por el mercado es lo que predominó, es decir la evaluación de impacto; en este caso la pregunta es ¿cómo se sabe que el efecto se debe al instrumento?

Figura 1: Proceso de evaluación de políticas



3 Nirenberg, O., Brawerman J, Ruiz V. "Evaluar para la transformación: innovaciones en la evaluación de programas y proyectos sociales". Paidós – Buenos Aires. 2000. Citado textualmente de Esteban Baltazar "Modelos de evaluación de políticas y programas sociales en Colombia" – 2008.

4 Se vincula con los métodos de "evaluación de resultados", "evaluación de impacto" y "experimental".

5 Se corresponde directamente con los métodos de "evaluación de impacto" y "experimental".

6 Ver la figura de abajo tomada de Najj y Astrand. "Outcome indicators for the evaluation of energy policy instruments and technical change – Energy Policy – mayo 2005.

7 Presidencias Reagan y Bush.

Para efectos de cumplir con lo solicitado en los términos de referencia de este proyecto “Propuesta de Instrumentos para facilitar medidas de eficiencia energética en el sector industrial de México”, se hizo una extensa búsqueda bibliográfica sobre el tema, con lo que se llegó a la conclusión de que no se desarrollarían propuestas que permitan implementar una evaluación ex - ante de las políticas y estrategias planteadas, sino a una evaluación basada, en muchos casos, en lo que se denomina “Theory-based-policy evaluation”⁸.

De dicha teoría pueden extraerse conclusiones valiosas, que, como se plantea anteriormente, se trata más de evaluaciones ex – post mediante indicadores, que modelos a priori para seleccionar políticas y estrategias. En todo caso las mismas surgirían de las experiencias fallidas del pasado, utilizando algunos de los métodos incorporados en la tabla anterior.

Lo cierto es que, en muchos casos, las conclusiones de tales evaluaciones, no se vinculan, directamente, con la eficacia de los instrumentos para alcanzar los resultados esperados, sino con el proceso de implementación, la burocracia en la fijación de etapas, las falencias en las estrategias de coordinación y la ausencia de una adecuada articulación al interior de un plan con sinergias y convergencias entre los diferentes instrumentos.

Así las condiciones de éxito identificadas se asocian con:

- Existencia de objetivos y mandatos claros para la organización que los implementa.
- Definición de instrumentos donde exista habilidad para combinar flexibilidad y continuidad.
- El involucramiento de los actores sujetos y objeto de la política.
- La habilidad para adaptar e integrar políticas paralelas o existentes y desarrollar paquetes consistentes de políticas.
- La coordinación institucional plena que involucre a todos los actores gubernamentales que serían necesarios para el éxito del instrumento.

⁸ Uno de los ejemplos es el documento: “Theory-based policy evaluation of 20 energy efficiency instruments” Mirjam Harmelink; Lars Nilsson; Robert Harmsen – Energy Policy – mayo 2008.

2. La Evaluación de Impacto (de una medida de política)

Por tanto, en una evaluación de impacto no solamente queremos saber si un problema mejora o empeora, sino si la intervención pública ha tenido algo que ver. Se trata, en resumen, de establecer si puede atribuirse o no (y en qué medida) la causa del cambio en el problema a la intervención pública. Hacerlo de forma convincente, como veremos, es una tarea laboriosa.

Por consiguiente, evaluar el impacto de un programa implica ser capaz de aislar el efecto del programa en relación con todos estos otros fenómenos que afectan al problema o situación que la intervención pública pretende abordar.

Esta constatación nos lleva introducir lo que parece un pequeño matiz, pero que tiene en realidad una importancia crucial en la evaluación de impacto (y que, como veremos más adelante, es la principal fuente de quebraderos de cabeza metodológicos): la pregunta que la evaluación de impacto debe responder no es qué ha pasado después de poner en marcha una intervención pública (muchas cosas pueden haber influido), sino qué ha pasado en comparación con lo que habría ocurrido si la intervención no se hubiera llevado a cabo. Lógicamente, la diferencia entre lo que ha sucedido con el programa y lo que habría sucedido sin el programa puede atribuirse sola y únicamente al programa o, dicho de otro modo, la diferencia ha sido causada por el programa. Y esto es, precisamente, lo que busca la evaluación de impacto: lo que el programa ha causado, y no lo que ha sucedido al mismo tiempo que el programa.

Una de las reglas de oro presentes en casi todos los manuales de estadística es no confundir asociación con causalidad. Pero desenredar causalidad y asociación en el campo de las políticas públicas no siempre es tan sencillo.

Cuando observamos una asociación (por ejemplo, que participar en un programa está asociado a una mejora en un determinado resultado), es importante tener siempre presente que la causalidad es una explicación posible, pero no la única. El reto de la evaluación de impacto es, justamente, descartar explicaciones alternativas para poder atribuir, de la forma más convincente posible, la causalidad del cambio observado a la intervención pública.

$$IMPACTO = Y1 - Y0$$

Donde:

- Y1 son los resultados que han ocurrido con la intervención pública.
- Y0 son los resultados que se habrían dado en ausencia de la intervención pública, que de forma más técnica (y más breve) se denominan contrafactual.