



Facultad
de Ciencias
Económicas y
Empresariales

Departamento
de Economía
Aplicada y
Estadística



**Revista de Evaluación de
Programas y Políticas Públicas**
JOURNAL OF PUBLIC PROGRAMS AND POLICY EVALUATION

Trascendiendo la instrumentalidad: metodología y evaluación

Beyond instrumentality: Methodology and evaluation

Núm. 7 (2016), pp. 1-27

Gris-Legorreta, Perla Carolina*

Recibido: **enero, 2016**

Aceptado: **octubre, 2016**

JEL Clasif: R38

DOI: [10.5944/reppp.7.2016.15870](https://doi.org/10.5944/reppp.7.2016.15870)

* Perla Carolina Gris-Legorreta. Maestra en Administración y Políticas Públicas (CIDE) y Doctora en Política (Universidad de Sheffield, Reino Unido). Actualmente se desempeña como Directora General Adjunta de Evaluación en la Coordinación Nacional de PROSPERA Programa de Inclusión Social (México). Sus líneas de investigación son: evaluación de políticas públicas, rendición de cuentas, gobernanza. E-mail: xlalegorreta@hotmail.com.

Resumen

La evaluación de políticas públicas es un instrumento indispensable para la gobernanza debido a la creciente complejidad del contexto en el que se desarrolla la acción pública. Elegir el método para evaluar las políticas y programas públicos representa una etapa fundamental en los procesos de gobierno y puede influir en los resultados del análisis. Este artículo explora algunas dimensiones que deben considerarse y explicitarse al momento de elegir los métodos de evaluación. Con base en la definición, propósito y audiencia de la evaluación (qué, para qué y para quién se evalúa) se exploran los criterios que derivan en la decisión técnica acerca de cómo evaluar las políticas públicas.

Palabras clave: gobierno; evaluación; políticas públicas

Abstract

Policy evaluation represents a difficult task due to the complexity of the context in which it takes place, as well as for the dynamism of the public sector. The methodological choice for assessing public policies and programs represents a decision of great importance as it might influence the results of the analysis. This article explores some of the multiple aspects to be considered regarding the methodology for policy evaluation, it reflects over issues such as the object and audience of evaluation and the nuances of the technical choice of how to perform the assessment with objectivity and scientific rigor.

Key Words: government; evaluation; public policy

1. Introducción

La presente investigación ofrece argumentos teóricos y empíricos para establecer un marco de referencia que facilite la selección de los métodos para la evaluación de las políticas públicas. La agenda de investigación y la práctica administrativa enfocada en la evaluación no han profundizado lo suficiente acerca de los criterios asociados a este proceso los cuales influyen de manera importante en la selección del método para la evaluación. En este contexto, surge como una pregunta de investigación relevante: ¿Cuáles son los criterios que deben considerarse en la selección de los métodos de evaluación?

Para sustentar estas ideas, el artículo se divide en cinco apartados. La primera parte discute los propósitos de la evaluación, por qué y para qué se evalúa, como un primer acercamiento a las necesidades metodológicas de la evaluación. La segunda sección aborda el tema de la audiencia, es decir, para quién se evalúa y su vinculación con la selección del método. La tercera parte está dedicada al objeto de la evaluación, qué es lo que se evalúa y cuál es la relevancia de delimitar la unidad de análisis. El cuarto apartado se enfoca en modelos y las metodologías de evaluación. Finalmente, se presentan algunas reflexiones acerca de la política de evaluación implementada por el Gobierno Federal mexicano.

En las sociedades democráticas, el ejercicio del poder público requiere ser escrutado permanentemente. Se requieren mecanismos para que el ciudadano verifique que el poder conferido a las instituciones por la vía democrática es ejecutado conforme a sus demandas. La evolución de la administración pública hacia modelos como la nueva gestión pública y, más recientemente, la gobernanza (Bryson, Crosby & Bloomberg, 2014) reconoce la necesidad de establecer instrumentos de cooperación y corresponsabilidad entre autoridades y ciudadanos (Uvalle, 2011).

Dado que los problemas públicos crecen cuantitativamente y se tornan cada vez más complejos, los gobiernos nacionales y subnacionales han encontrado en las políticas públicas una herramienta de gestión útil para encauzar los recursos y esfuerzos del sector público hacia la atención de estos problemas (Jung, 2014). La existencia de múltiples actores con intereses diversos en un contexto de incertidumbre demanda el estudio y análisis de la interacción que se produce entre ellos (Weible *et al.*, 2012).

Una de las etapas del proceso de hechura de políticas públicas que ha adquirido mayor notoriedad en los últimos lustros es la evaluación, una herramienta dirigida a obtener información acerca de la gestión pública, la cual implica:

«reestablecer la confianza en la capacidad del gobierno para producir mejores servicios públicos, así como contar con un aparato gubernamental más inteligente que pueda reducir la brecha entre las expectativas ciudadanas y la capacidad gubernamental» (Santiso, 2015, p. 123).

La evaluación tiene como uno de sus principales objetivos aportar elementos que fortalezcan las etapas del proceso de políticas públicas, desde el planteamiento del problema y su incorporación a la agenda pública hasta su implementación y posible terminación. Posee también un rol importante en el fortalecimiento de la democracia a partir de su relación con la rendición de cuentas (Hanberger, 2006) y su contribución a la toma de decisiones (Watson, 2014).

Aun cuando esta actividad permite «identificar estándares adecuados para evaluar y utilizar métodos de análisis válidos y justos» (Picciotto, 2005, p. 134) y se desarrolla a través de instrumentos científicos y entendidos como neutrales, la evaluación es ineludiblemente política, un aspecto que no puede ignorarse (Palumbo, 1987).

A pesar de su creciente importancia, poco se ha reflexionado con respecto al proceso de evaluación y sus implicaciones. Preguntas como: ¿Quién elige qué evaluar? ¿Bajo qué criterios se definen las metodologías de evaluación? ¿Qué uso se dará a la información producida? ¿Quién es el público objetivo de la información? ¿Quién realizará las evaluaciones y quién será el encargado de validar el trabajo del evaluador? son interrogantes que demandan mayor atención.

El presente documento no pretende responder todas las interrogantes. Se reflexionará únicamente con respecto al elemento que determina que éste sea un ejercicio útil y, sobre todo, válido: la metodología. En el método y en los estándares seleccionados recae gran parte de la validez debido a que explicitan el mecanismo mediante el cual se obtienen los hallazgos. Elegir la metodología conlleva el análisis de diversos elementos que trascienden la solidez técnica de un instrumento en particular, existen valores, principios y posiciones políticas involucrados que no pueden ignorarse (Greene *et al.*, 2007). Este artículo desarrolla un marco de referencia para la selección de la metodología de evaluación con base en el análisis de variables identificadas en la literatura especializada que se consideran relevantes para esta decisión (Figura 1).

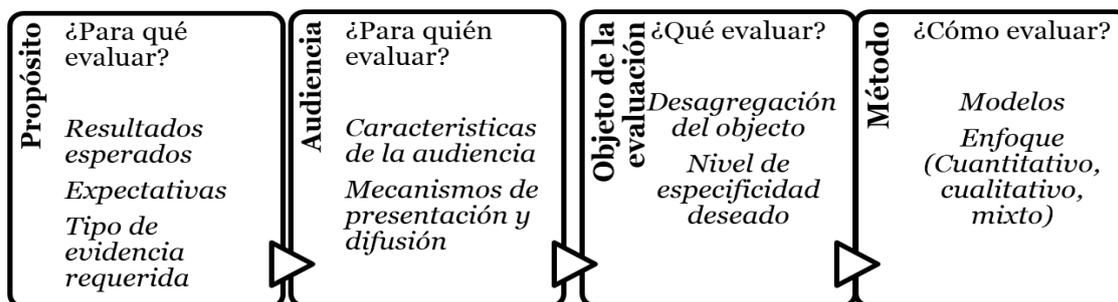


Figura 1: Marco de Referencia para la selección metodológica de la evaluación.

2. Evaluar, ¿para qué? Los propósitos de la evaluación de políticas públicas

Como un mecanismo de apoyo en la toma de decisiones (Campbell *et al.*, 2009), la evaluación requiere que su finalidad sea clara tanto para quienes la llevan a cabo como para a quien está dirigida, de lo contrario puede transformarse en una actividad sin sentido dentro del quehacer gubernamental y percibirse como un ejercicio estéril (Feinstein, 2002). El uso de la evaluación está fuertemente vinculado con la capacidad de las organizaciones de asimilar y entender el conocimiento (Cousins *et al.*, 2014). Esto exige cuestionar las razones que justifican su existencia, es decir, su propósito: ¿evaluar con qué fin? y ¿qué es lo que se espera obtener de la evaluación?

Existen diversas motivaciones para llevar a cabo un escrutinio de la acción pública. De acuerdo con Chelimsky (1987), la evaluación es una actividad con múltiples fines: 1) identifica los resultados de las políticas y programas, así como problemas públicos y posibles alternativas de solución, 2) fortalece el ejercicio público a partir de estudiar los procesos administrativos, el funcionamiento de las organizaciones y la vinculación gubernamental con otros agentes de la esfera pública y 3) apoya las reformas gubernamentales.

A través de proveer información técnica y objetiva acerca del desempeño del ámbito público, la evaluación brinda elementos para que los tomadores de decisiones actúen de manera racional (Hertting & Vedung, 2012). Asimismo, busca identificar una relación causal entre una política pública y sus efectos, es decir, «el impacto o beneficio neto de un programa u organización» (Wholey, 2001, p. 345) y responde a la necesidad de mecanismos de medición y análisis de políticas y programas. Por lo tanto, gran parte de las actividades de monitoreo y evaluación recaen en el ámbito de la eficiencia y el impacto pues permiten comprender por qué y en qué condiciones una política funciona (Sanderson, 2001). La identificación de efectos netos atribuibles a un determinado programa —aislando la influencia de otras posibles variables— demanda la utilización de herramientas metodológicas específicas que analicen el efecto de un programa con base en un grupo de tratamiento y uno de control (contrafactual) (Gertler *et al.*, 2011).

En el marco de la gobernanza¹, los propósitos de la evaluación son fundamentalmente el desarrollo y la rendición de cuentas (Hertting & Vedung, 2012). El desarrollo «(mejora, aprendizaje, buenas prácticas) puede ser formativo² aspira a orientar la modificación del objeto evaluado en términos de su implementación» (Hertting & Vedung, 2012, p. 37). El gran propósito *instrumental* de la evaluación es, por lo tanto, determinar no solo qué políticas funcionan sino por qué funcionan, qué las hace funcionar y cómo pueden mejorarse.

Las evaluaciones de procesos³, por ejemplo, permiten señalar aspectos específicos que obstaculizan la operación y que pueden modificarse para obtener mejores resultados.⁴ El estudio detallado de la operación requiere desarticular los procesos involucra-

1 En el presente documento, gobernanza se entiende como «un cambio en el significado de gobierno, el cual alude a un nuevo proceso de gobernar [...] a la auto-organización, redes inter-organizacionales interdependientes, intercambio de recursos, reglas del juego y una autonomía significativa del Estado.» Rhodes, R. A. W. (1997). *Understanding governance: Policy networks, governance, reflexivity and accountability*. Buckingham: Open University Press. P. 15.

2 El concepto *formative evaluation* se refiere a un tipo de evaluación cuyo objetivo principal «es obtener información durante el proceso de implementación que permita valorar el desarrollo del programa» en contraposición al concepto de *summative evaluation* que «busca obtener información acerca de la efectividad de un programa, tras su implementación.» Bennett, J. (2003). *Evaluation methods in research*. Londres: Continuum. Pp. 9-10.

3 De acuerdo con el modelo emitido por el CONEVAL para el desarrollo de la evaluación de procesos, el objetivo de este instrumento es «realizar un análisis sistemático de la gestión operativa del programa que permita valorar si dicha gestión cumple con lo necesario para el logro de metas y objetivo del programa. Así como, hacer recomendaciones que permitan la instrumentación de mejoras.» CONEVAL (2013). *Modelo de Términos de Referencia para la Evaluación de Procesos*. P. 1.

4 En 2012, por ejemplo, fueron evaluados siete procesos del programa de la Secretaría de Salud «Protección contra Riesgos Sanitarios», con base en este análisis se presentaron distintas recomendaciones dirigidas a mejorar la operación, por ejemplo, «fortalecer la gestión de Recursos Humanos.» Secretaría de Salud. (2013). *Protección contra Riesgos Sanitarios. Evaluación de Procesos*. P. 219.

dos y analizar detalladamente la sucesión de actividades dirigidas a un fin en particular. Cuando el propósito de una evaluación es de carácter instrumental, el método a utilizar debe permitir un nivel de especificidad que genere hallazgos de utilidad. Por ejemplo, en el marco de las evaluaciones de procesos, es un método recurrente realizar entrevistas a los operadores para entender el funcionamiento de un programa (Consejo de Evaluación—Universidad de Chile, 2011).

La discusión acerca del propósito de la evaluación se ha hecho cada vez más sofisticada a partir de su vinculación con conceptos como la gobernanza y la rendición de cuentas.⁵ En el ámbito de la gobernanza, la evaluación es vista como un instrumento útil dado que existe un número mayor de actores participando en el proceso de políticas públicas que deben coordinarse (Hanberger, 2006). La rendición de cuentas, por su parte, ha posicionado a la evaluación como un elemento indispensable para atender las demandas de información acerca del desempeño gubernamental pues permite obtener «evidencia para que los responsables de la acción gubernamental puedan cumplir con la obligación de informar de manera permanente acerca de sus decisiones y acciones» (Gris-Legorreta, 2011, p. 203). Las evaluaciones con este propósito «producen información valorativa para permitir la toma de decisiones acerca de la continuidad, expansión, reducción o eliminación de un programa» (Widmer & Neuenschwander, 2004, p. 391); sin embargo, demandan algunas condiciones:

«Los sistemas de evaluación y rendición de cuentas tienen que ser consistentes con los roles de los funcionarios y sus circunstancias, de tal forma que su retroalimentación favorezca un comportamiento administrativo y un proceso de toma de decisiones apropiado para cada situación» (Zapico-Goñi, 2007, p. 437).

La relación entre ambos conceptos es recíproca. La evaluación favorece, por una parte, la generación de evidencia que los servidores públicos pueden utilizar para responder los cuestionamientos de un tercero. La rendición de cuentas, por su parte, al llevar implícito un elemento punitivo, es el mecanismo ideal para que la evaluación no constituya un esfuerzo estéril. El incumplimiento de objetivos, el uso indebido de los recursos, la planeación deficiente, entre otros aspectos, es posible evidenciarlos a través de la evaluación para que en el marco de un sistema efectivo de rendición de cuentas se puedan corregir las fallas y sancionar a los responsables. El concepto de *accountability evaluation* engloba esta idea al señalar como su propósito:

«Indagar acerca de cómo los agentes han ejercido el poder que les fue delegado de tal manera que su trabajo pueda ser juzgado por sus superiores y tomar acciones al respecto [...] es una herramienta para que los superiores evalúen a sus subordinados y que éstos se hagan responsables de sus acciones» (Hertting & Vedung, 2012, p. 38).

Independientemente de que el objeto y la metodología de evaluación sean los adecuados, promover la utilización de la información derivada de las evaluaciones representa otro desafío para los involucrados en el diseño y la planeación de la evaluación en

⁵ Rendición de cuentas se entiende como «la relación entre un actor y una audiencia en la cual el actor tiene una obligación de explicar y justificar su conducta, mientras que la audiencia puede inquirir y juzgar, y el actor puede enfrentar consecuencias.» Bovens, M. (2007). *Analysing and Assessing Accountability: A Conceptual Framework*. *European Law Journal*, 13 (4), 447-468. P. 450.

términos del «uso efectivo y potencial» (Feinstein, 2002, p. 434). El concepto de *utilization-focused evaluation*⁶ conlleva la adopción de ciertas medidas dentro del desarrollo de una evaluación para incentivar su uso. Un ejemplo es el involucramiento de todos los interesados (*stakeholders*) en el proceso y que estos reciban capacitación para potenciar su contribución (Patton, 1986, p. 335). En este sentido, Leviton & Hughes (1981) identifican cinco grupos de variables vinculadas con el uso:

«1) La relevancia de la evaluación para las necesidades de los usuarios potenciales; 2) El alcance de la comunicación entre usuarios potenciales y quienes generan la evaluación; 3) La transformación de la evaluación en implicaciones de política; 4) La credibilidad y confianza en las evaluaciones y 5) El compromiso y el apoyo de los usuarios [...]» (Leviton & Hughes, 1981, p. 525).

La atención a estas variables puede contribuir a que la utilidad no solo dependa de la voluntad política de los tomadores de decisiones lo que permitiría transitar de un enfoque en el uso hacia uno que valorara la influencia de la evaluación (Henry & Mark, 2003). Esta reflexión acerca del propósito de la evaluación permite indagar en aspectos más específicos como es el modelo de evaluación.⁷ Hansen (2005) observa que existe un vínculo importante entre estas dos variables. Si el propósito de la evaluación es establecer un mecanismo de control, existirá entonces una mayor preocupación por generar evidencia acerca del nivel de cumplimiento de los objetivos. Si, por el contrario, el propósito está vinculado con el aprendizaje, serán el diálogo y la deliberación entre participantes (Tabla 1).

Tabla 1. Selección del modelo de evaluación con base en su propósito.

Diseño de la evaluación	Propósito de la evaluación	
	Control	Aprendizaje
Modelo de evaluación	Énfasis en los resultados	Énfasis en el proceso
Organización	Top-down	Auto-evaluación , evaluación independiente de pares
Criterio	Derivado de los objetivos	Derivado de las solicitudes de los stakeholders
Método	Principal énfasis en la medición cuantitativa de los efectos de la política	Principal énfasis en el análisis cualitativo de procesos
Difusión de los resultados	De naturaleza jerárquica	A través de procesos interactivos con distintos grupos de interés
Uso	Mecanismos de sanción y control	Auto-reconocimiento y desarrollo organizacional

Fuente: Hansen (2005).

6 Al respecto, Patton (1986) ha sido uno de los principales promotores del concepto de *utilization-focused evaluation*, el cual consiste fundamentalmente en promover la utilidad de la evaluación, es decir, evitar que se convierta en una actividad estéril y sin efectos sobre el proceso de políticas públicas.

7 Un modelo de evaluación «estipula la pregunta que una evaluación busca responder, al mismo tiempo que especifica cómo establecer los criterios para la valoración» (Hansen, 2005, p. 448). Así, la elección de un modelo de evaluación acota el universo de metodologías que pueden utilizarse. Existe una vasta literatura acerca de modelos de evaluación y del diseño de la evaluación.

Los modelos de evaluación pueden clasificarse entonces a partir de la naturaleza de la información que se busca generar y de los resultados esperados. En este sentido, Owen & Rogers (1999) desarrollaron una tipología de modelos⁸ que permite identificar las preguntas que deben plantearse como parte del diseño de la evaluación y de la elección de la metodología. Este modelo provee información importante acerca del propósito y las expectativas de la evaluación. La selección del modelo de evaluación, por tanto, no debe estar constreñida a aquéllos que favorecen el uso, sino debe mantenerse una posición abierta a la valoración de otras variables (Contandriopoulos & Brousselle, 2012).

3. Evaluar, ¿para quién? El público de las evaluaciones

Uno de los propósitos de la evaluación es que sus resultados sean utilizados en la mejora de las políticas públicas. Sin embargo, este uso está en función de si son relevantes para alguien, esto es, quién considera útil esa información, a quién le sirve y por qué.

Este aspecto puede verse desde la óptica económica. Existen por un lado, *productores de evaluaciones* (*think-tanks*, universidades, organismos internacionales, consultorías, agencias gubernamentales, entre otros) y por otro, *consumidores de evaluaciones* (academia, medios de comunicación, tomadores de decisiones, *policymakers*, sociedad civil, etc.), es decir, una oferta y demanda de evaluaciones. La importancia radica esencialmente en vincular ambas partes para promover la utilidad de la evaluación. Esto implica identificar claramente a la audiencia de la evaluación, ¿a quién se presentarán los hallazgos de la evaluación? ¿qué utilidad tienen estos para esa audiencia?

Chelimsky (1987) identificó que existen distintos usuarios de la evaluación en el ámbito público. En el Poder Ejecutivo, son identificables los altos mandos de las agencias gubernamentales, los operadores de los programas, así como las autoridades en materia presupuestaria y administrativa. En el Poder Legislativo los usuarios son las distintas comisiones legislativas así como las instancias encargadas de la fiscalización (Palumbo, 1987, p. 76). Existen también otros usuarios tales como los medios de comunicación, los académicos y la sociedad civil, por mencionar algunos (Weiss, 1998b).

Una de las audiencias más visibles de la evaluación son los *policymakers*. La información generada por las evaluaciones facilita a esta audiencia el proceso de toma de decisiones en un contexto incierto, cambiante y de fuertes presiones políticas. Estos actores valoran el uso de esta información:

- 1) Porque les ayudan a tomar buenas decisiones;
- 2) Porque les evita tomar malas decisiones;
- 3) Porque les ayuda a ganar el respeto de sus colegas y 4) Porque les permite obtener mayor apoyo en el ámbito legislativo (Bogenschneider & Corbett, 2010, pp. 27-31).

Debe existir, por tanto, una necesidad de evidencia y esta poseer un determinado nivel de relevancia para una organización (Neuman *et al.*, 2013).

8 En el apartado 4 del presente documento se profundiza en la tipología de modelos de evaluación desarrollada por Owen & Rogers (1999).

Preguntarse evaluar, ¿para quién? se enlaza con lo ya discutido en la sección anterior acerca de los factores que favorecen el uso de las evaluaciones. Reflexionar acerca de la audiencia de la evaluación es un factor indispensable en el diseño debido a que influye de manera considerable en el uso potencial que se dará a los hallazgos. Por consiguiente, conocer a fondo las características de la audiencia principal es necesario para contar con elementos para la toma de decisiones, por ejemplo, acerca del método y la estrategia de divulgación de los hallazgos.

Las expectativas que tienen sobre una evaluación los distintos actores son de distinta naturaleza y la información generada por una evaluación variará con base en las decisiones tomadas desde su diseño. Como lo ejemplifica Weiss, «los *policymakers* se preocupan de los efectos netos de un programa, mientras que los operadores del programa están más interesados en los efectos diferenciados respecto de otras alternativas» (Weiss, 1998b, p. 29).

La identificación del propósito y de los actores involucrados sienta las premisas para el diseño de la evaluación pues permiten establecer con mayor certeza el tipo de información que se requiere, así como el uso potencial que se le dará a la misma. Un aspecto interesante lo constituye el Poder Legislativo como una de las principales audiencias de la evaluación. De acuerdo con la investigación de Vanlandingham (2011) en los Estados Unidos, el valor otorgado a los productos de la evaluación por los usuarios se incrementa en tanto que exista una mayor interacción entre los evaluadores y los actores involucrados en este proceso. De ahí que conocer las características de la audiencia potencial sea un elemento sustantivo a considerar dentro de la selección metodológica.

La audiencia influye también en cómo se diseminan los resultados. La información que es relevante para un grupo de tomadores de decisiones no puede tener las mismas características que otra dirigida a la sociedad civil (Lawrenz *et al.*, 2007). Retomando el caso de los *policymakers*, estos tienen la alternativa de buscar y utilizar evidencia proveniente de distintas fuentes, incluyendo la evaluación (Lundin & Öberg, 2014). La información que consideran de interés tiene características muy particulares, como por ejemplo, que sea objetiva y oportuna; que tenga implicaciones de política; que las recomendaciones puedan implementarse a un costo razonable; que los resultados se presenten de manera clara y concisa; que exista rigor técnico, entre otras (Bogenschneider & Corbett, 2010).

Las características de la audiencia demandan que las evaluaciones cumplan con criterios determinados en contenido, presentación y difusión de los resultados. Ignorar las necesidades de dicha audiencia es prácticamente condenar una evaluación al archivo muerto, sin considerar el *para quién*, el *por qué* y el *cómo* pierden gran parte de su sentido (Gris-Legorreta, 2013). La identificación y comprensión de la audiencia son fundamentales para transitar de la descripción de la evidencia a la argumentación (Majone, 2005).

4. Evaluar, ¿qué? La evaluación de políticas públicas y su objeto

Otro de los elementos a considerar en el proceso de evaluación es la identificación del objeto de evaluación. Dicho objeto puede ser no únicamente una política o un programa sino también una organización, un proceso e inclusive una persona. Este representa la unidad de análisis que determinará muchas de las decisiones relativas a la evaluación; sus características y naturaleza deben considerarse para coadyuvar a que el proceso de evaluación y sus resultados cumplan con las expectativas. Vartiainen (2002) reflexionó al respecto retomando algunas discusiones presentes en la literatura:

«De acuerdo con Argyris (1962), en el contexto de un proceso de evaluación, se debe decidir si se evaluará la acción individual o la acción colectiva, Pennings & Goodman (1979) y Manns & March (1978) sugieren que una evaluación debe concentrarse en la coordinación y el cumplimiento de los objetivos de distintas unidades de una organización; mientras que Weiss reflexiona acerca de si las organizaciones deben evaluarse como entidades o únicamente con base en los programas que operan ...» (Vartiainen, 2002, p. 362).

No es trivial que la acción pública traducida en políticas, programas, estrategias, acciones, etc., genera la complejidad de diferenciar entre lo que es relevante evaluar y lo que no. Para Owen & Rogers (1999) el objeto puede clasificarse en cinco categorías: *programas, políticas, organizaciones, productos e individuos*. Esta clasificación no es limitativa pues es posible desagregar cada una de las categorías para tener una idea más clara del objeto de la evaluación y llegar hasta el nivel de especificación que cumpla con las expectativas. Por ejemplo, la evaluación no puede utilizarse sólo para identificar y cuantificar los efectos de un programa de vacunación sobre la población infantil, sino también para aspectos más específicos como la eficiencia de las instituciones de salud para distribuir las vacunas.

Las características del objeto inciden en el tipo de evaluación que puede llevarse a cabo. Esto constituye uno de los riesgos más importantes a los que está expuesta una política de evaluación. La homologación de la evaluación, a partir de la aplicación de un único instrumento metodológico para un conjunto de programas conduce a la obtención de resultados irregulares. Mientras que una determinada metodología se ajusta con mayor facilidad a las características de un programa, la misma puede resultar deficiente para otros por distintos motivos. En primer lugar, porque las demandas de información y contenido que requiere un determinado instrumento no son satisfechas por todos los programas. En segundo, porque aun cuando los resultados a obtener serán dispares, unos de más calidad que otros, el proceso de evaluación constituye para todos los programas (sin importar lo relevantes u oportunos que serán los hallazgos) un esfuerzo de gestión y organización que en muchos casos resultará estéril. Finalmente, porque la idea de que *one size fits all* aunque facilita el proceso de evaluación y la obtención de información estandarizada, ha tenido costos en términos de calidad y utilidad (Cejudo, 2011).

Aspectos como el propósito y el objeto de evaluación determinarán el tipo de evidencia a obtener (Berriet-Sollicet *et al.*, 2014).⁹ La naturaleza de la evidencia requiere un tratamiento metodológico específico que debe definirse en la etapa del diseño de la evaluación.

Esto conduce a la distinción entre productos y resultados, o como señala la literatura anglosajona, *outputs* y *outcomes*. Un *output* puede entenderse como «los bienes o servicios (generalmente estos últimos) que las agencias gubernamentales proveen a los ciudadanos, empresas o a otras agencias gubernamentales», mientras que los *outcomes* se definen como «los impactos o las consecuencias que tienen sobre la comunidad, los outputs o actividades gubernamentales» (Kromann *et al.*, 2001, p.10). Esto último puede entenderse también como «lo que se prevé que los participantes en un programa sean capaces de hacer al final de su participación» (Patton, 1986, p. 96). En materia de evaluación esta distinción determina no sólo el tipo de método a elegir, sino la naturaleza de los resultados a obtener e inclusive el efecto que estos tendrán sobre todo el proceso de políticas públicas. La valoración de los *outcomes* es necesaria como parte de la ecuación para entender el proceso de políticas, en específico, para saber cómo la implementación de un conjunto de acciones produce un efecto dentro del ámbito de la gobernanza (Boardman, 2014).

Esta diferenciación ha conducido también a la distinguir entre monitoreo y evaluación, lo cual es relevante entender con mayor detenimiento dado que el desarrollo de sistemas de evaluación generalmente incorpora un componente dirigido al monitoreo de la actividad gubernamental. Como resultado, estas actividades se identifican como complementarias. Mientras que el monitoreo implica la obtención sistemática de información acerca del cumplimiento de objetivos con base en indicadores; la evaluación está dirigida a encontrar vínculos causales entre una intervención y los resultados obtenidos, así como a entender por qué se obtuvieron o no los resultados esperados (Kusek & Rist, 2004). Autores como Talbot (2010) han reflexionado acerca de las diferencias entre monitoreo y evaluación (Tabla 2).

Tabla 2. Principales diferencias entre evaluación y monitoreo.

Variable	Evaluación	Monitoreo
Objeto	Políticas y programas	Organizaciones
Frecuencia	Episódica	Periódica
Métodos	Mixtos (cuantitativos y cualitativos)	Principalmente cuantitativos
Ejecutor	Fundamentalmente realizada o comisionada por agencias o departamentos de política pública	Fundamentalmente ejecutado o comisionado por las áreas financieras u operativas

Fuente: Talbot (2000).

⁹ Estos autores identifican tres tipos de evidencia: «evidencia de hechos (identificación de elementos existentes en un contexto, por ejemplo, especies de aves en un lugar determinado); evidencia acerca de un efecto/diferencia producida (identificación de elementos que demuestren el beneficio o daño generado por una intervención) y evidencia acerca del mecanismo de un fenómeno (identificación de elementos acerca de cómo un aspecto tiene efectos sobre otro).» Berriet-Sollicet, M., Labarthe, P. & Laurent, C. (2014). Goals of evaluation and types of evidence. *Evaluation*, 20 (2), 195-213. P. 198.

Esta reflexión es útil para comprender cómo los distintos objetos de la evaluación conducen a hallazgos diferentes y demandan metodologías diversas. Pareciera entonces que los esfuerzos de una política de evaluación deben enfocarse en la identificación y cuantificación de los efectos de una política sobre la población a la que está dirigida, mientras que el monitoreo hacia la medición de la eficiencia en la generación y entrega de productos y servicios públicos. Este aspecto no es menor pues los requerimientos técnicos y administrativos de estas actividades son distintos por razones como la periodicidad, el nivel de profesionalización de los ejecutores, los recursos financieros y materiales, entre otros. La capacidad técnica en materia de evaluación, por ejemplo, constituye un elemento determinante para el establecimiento efectivo de un sistema de evaluación (Gaarder & Briceño, 2010, p. 20).

En consecuencia, la identificación del objeto de evaluación constituye un elemento importante del diseño de la evaluación al acotar la unidad de análisis alrededor de la cual se seleccionará el modelo y, posteriormente, la metodología a seguir con base en los tres aspectos señalados: la especificidad del objeto, su naturaleza y características y el tipo de resultados que se espera obtener.

5. Evaluar, ¿cómo? La metodología de la evaluación de políticas públicas

La selección metodológica responde a un variado número de factores vinculados con la naturaleza y las características del objeto de evaluación pero también con los resultados que se esperan de la misma. El desarrollo de los distintos modelos de evaluación ha incorporado elementos como el objeto y el propósito de la evaluación para robustecer el diseño de la valoración de las políticas. De acuerdo con Owen & Rogers (1999) existen cinco modelos generales de evaluación: *proactivo*, *clarificativo*, *interactivo*, *monitoreo y de impacto*, basados en el planteamiento de distintas preguntas para valorar diversos aspectos del objeto de evaluación (Tabla 3).

Tabla 3. Tipología de modelos de evaluación.

Modelo	Propósito	Preguntas que busca responder	Principales enfoques
Evaluación proactiva	Fortalecer los procesos de toma de decisión previo a la creación de un programa o a la implementación de una política para la atención de un problema público.	<p>¿Existe la necesidad de crear un programa?</p> <p>¿Qué áreas del problema identificado se prevé atender mediante el programa?</p> <p>¿Se han identificado alternativas de solución al problema identificado?</p> <p>¿Qué sugiere la literatura acerca de posibles soluciones?</p> <p>¿Qué conocimiento se tiene del problema que pretende atender el programa?</p>	<ul style="list-style-type: none"> • Análisis de necesidades (<i>needs assessment</i>) • Revisión de literatura • Revisión de mejores prácticas
Evaluación clarificativa	Promover un mejor entendimiento de la estructura interna y el funcionamiento de una política o programa.	<p>¿Cuáles son los resultados esperados del programa y cómo se prevé lograr dichos resultados?</p> <p>¿Cuál es la lógica detrás del programa?</p> <p>¿Qué elementos del programa requieren ser modificadas para maximizar el potencial del programa y lograr los objetivos planteados?</p> <p>¿Es plausible el programa?</p> <p>¿Qué aspectos del programa son susceptibles de ser monitoreados o de evaluar su impacto?</p>	<ul style="list-style-type: none"> • Lógica del programa (<i>programa logic development</i>) • Certificación¹⁰ (<i>accreditation</i>)
Evaluación interactiva	Este modelo de evaluación provee información acerca de la implementación de un programa o acerca de alguno de sus componentes o actividades.	<p>¿Qué se busca lograr a través del programa?</p> <p>¿Cómo se presta el servicio?</p> <p>¿Funciona la prestación del servicio adecuadamente?</p> <p>¿Es consistente con el diseño del programa?</p> <p>¿Cómo podría ser más efectiva la prestación del servicio?</p>	<ul style="list-style-type: none"> • Evaluación responsiva • <i>Action research</i> • Control de calidad • <i>Empowerment evaluation</i>

(Continúa)

10 Se refiere fundamentalmente a realizar una evaluación con base en la determinación de si una agencia puede o no llevar a cabo un programa en un período de tiempo determinado.

Modelo	Propósito	Preguntas que busca responder	Principales enfoques
Monitoreo	Identificar el progreso de un programa en curso, por lo general para verificar el cumplimiento de los objetivos y su vinculación con el gasto.	¿El programa está atendiendo a la población objetivo? ¿Existen diferencias en la implementación en los distintos lugares donde opera el programa? ¿Qué cambios han existido en la implementación durante distintos períodos de tiempo? ¿Los costos han aumentado o disminuido? ¿Qué ajustes podrían hacerse para aumentar la eficiencia del programa? ¿Qué ajustes podrían hacerse para aumentar la efectividad del programa?	<ul style="list-style-type: none"> • Valoración del desempeño • Análisis de componentes del programa a través de indicadores • Análisis de procesos
Evaluación de impacto	Evaluar el efecto que tiene un programa, se valoran los resultados obtenidos tras la implementación de un programa.	¿Se implementó el programa conforme a lo establecido? ¿Se alcanzaron los objetivos planteados? ¿Se atendieron las necesidades de la población objetivo? ¿La implementación del programa condujo a los resultados esperados? ¿Cómo afectaron las diferencias en la implementación los resultados del programa? ¿Cuál fue la relación costo-beneficio?	<ul style="list-style-type: none"> • Evaluación de objetivos • Evaluación de la relación entre proceso y resultados (<i>process-outcomes studies</i>) • Evaluación basada en las necesidades • <i>Goal-free evaluation</i> • Auditoría del desempeño

Fuente. Elaboración propia con base en Owen & Rogers (1999).

La metodología permite que la evaluación se realice a través de un procedimiento ordenado, sistemático y replicable, lo que facilita la estandarización e implementación de la evaluación (Chen, 1994). Efectivamente, cuando existe un dominio de una herramienta metodológica, el desarrollo de una evaluación se torna más fácil pues se conoce con mayor precisión el tipo de datos requeridos para su tratamiento; sin embargo, el contexto y las características de un programa son elementos importantes a considerar. Cuando la planeación de la evaluación se encuentra en la etapa de la selección de metodología, el análisis contextual y del objeto del programa debe haber sido ya agotado pues constituyen variables importantes para definir qué herramienta metodológica es más pertinente.

El tipo de herramientas para evaluar una política o un programa pueden ser de distinta naturaleza. Particularmente, la selección de la metodología se ha centrado en la elección entre métodos cuantitativos y cualitativos o, más recientemente, en la valoración de enfoques mixtos (Johnson & Onwuegbuzie, 2004). Las ventajas y desventajas de estos enfoques han sido ampliamente discutidas tanto en las ciencias sociales (Creswell, 1994; Matthews & Ross, 2010; Rossi *et al.*, 1979; Øyen, 1990) como en el campo de la evaluación (Bamberger *et al.*, 2010; Rao & Woolcock 2003). Patton (2002) ha analizado este tema a partir de una serie de preguntas clave previo a la selección del enfoque (Figura 2).

¿Cuál es el propósito?	<ul style="list-style-type: none"> • Investigar: Contribuir a la generación del conocimiento. • Evaluar: Mejorar un programa y apoyar la toma de decisiones. • Interés personal: Obtener información para uno mismo.
¿Quién es la audiencia principal de los hallazgos?	<ul style="list-style-type: none"> • Académicos, investigadores. • Servidores públicos, participantes del programa • Amigos, familia
¿Qué tipo de preguntas guiarán la investigación?	<ul style="list-style-type: none"> • Derivadas u orientadas por la teoría • De aplicación práctica • Derivadas de intereses personales
¿Qué tipo de datos se utilizarán?	<ul style="list-style-type: none"> • Cualitativos: Entrevistas, observación directa, documentos • Cuantitativos: Encuestas, experimentos, fuentes secundarias • Mixtos: ¿Qué tipo de combinación de datos? ¿Cuáles serán los predominantes?
¿Con qué recursos se cuenta?	<ul style="list-style-type: none"> • Financieros • Tiempo • Recursos humanos • Acceso, relaciones personales
¿Qué criterio se utilizará para juzgar la calidad de los hallazgos?	<ul style="list-style-type: none"> • Criterios metodológicos tradicionales: rigor, calidez, confiabilidad, generalización • Criterios de evaluación: utilidad, factibilidad, pertinencia • Criterios no tradicionales: Credibilidad del investigador, diversidad de perspectivas.

Figura 2: Preguntas guía para la selección del enfoque metodológico.

Fuente: Patton (2002).

Esta etapa implica tomar decisiones en materia de «la estrategia de evaluación, el diseño, la selección de la población y la muestra, la recolección de datos, así como del tipo de análisis a realizar» (Posavac & Carey, 2003, p. 40) por lo que el rigor técnico y científico de un método en particular no puede subsanar las posibles deficiencias que existan en otros ámbitos y mucho menos debe constituir el criterio principal de selección.

Las decisiones metodológicas no están sólo dirigidas a entender el funcionamiento o los efectos de una intervención, sino también aquellos casos en los que la intervención ha fracasado pues permite «predecir posibles fallas e incrementar la probabilidad de éxito» (Durand *et al.*, 2014, p. 405).

El uso potencial de la evaluación con base en su propósito, el objeto de análisis y la audiencia a la que está dirigida muestra la importancia de la decisión en función de que ésta no solo afectará el ámbito técnico de la recolección y análisis de la información, sino también el rigor. El cuestionamiento del rigor puede tener implicaciones importantes en materia de utilización si los hallazgos no son considerados evidencia suficiente de las conclusiones de un estudio (Braverman, 2012).

Como consecuencia de esto, la discusión acerca de la metodología se ha trasladado al tipo de evidencia que ésta puede producir: «el conflicto actual no radica en lo deseable que es generar evidencia o en la necesidad de considerar el valor relativo de distintas metodologías, el debate está en cambio fundamentalmente en cuándo o bajo qué circunstancias distintas metodologías producen la evidencia más útil» (Greene *et al.*, 2007, p. 120). En este punto, vuelven a resonar fuertemente las preguntas de para qué y para quién se evalúa. Existen tres categorías en las cuales puede clasificarse la evidencia acerca de la selección de un método para sustentar esta etapa del diseño de la evaluación (Figura 3).

<p>Evidencia basada en resultados¹¹</p> <p><i>Outcome-based evidence</i></p>	<p>Se relaciona con el impacto de la selección de un método determinado en términos de la exactitud y precisión de los hallazgos y la persuasión de los resultados.</p>
<p>Evidencia basada en experiencia</p> <p><i>Practice-Based Evidence</i></p>	<p>Se refiere a la evidencia resultante del conocimiento profesional acumulado acerca de qué métodos de evaluación han demostrado ser efectivos, así como el conocimiento acerca de las condiciones necesarias para su eficaz implementación.</p>
<p>Revisión crítica de metodologías</p>	<p>Sugiere valorar los distintos métodos a partir del análisis de aspectos como las implicaciones éticas de estos, la posible violación de derechos o si podrían producir efectos negativos sobre la población.</p>

Figura 3: Evidencia y selección del método.

Fuente: Elaboración propia con base en Greene *et al.* (2007).

Después de presentar estas discusiones acerca de la selección de la metodología, es oportuno exponer algunas de las herramientas metodológicas más relevantes identificadas en la literatura¹² para observar con mayor de detalle las características de cada una de estas y lo que ofrecen para la evaluación de políticas (Tabla 4).¹³

11 Para mayor referencia del tipo de modelos vinculados con esta categoría se sugiere ver: Shadish, W.R. & Rindskopf, D. M. (2007). Methods for evidence-based practice: Quantitative synthesis of single-subject designs. *New Directions for Evaluation* (113), 95-109.

12 Existe una gama muy amplia de herramientas para la evaluación de políticas por lo que la relación presentada en este artículo no es limitativa.

13 El propósito de esta tabla brindar al lector un panorama general de las principales herramientas metodológicas utilizadas para la evaluación de políticas públicas, por lo que la información que se presenta

Tabla 4. Herramientas metodológicas para la evaluación de políticas.

Enfoque	Herramienta	Descripción	Resultados esperados
Cuantitativo	Evaluación de impacto	Se utiliza con el propósito de evaluar si una intervención ha producido los efectos esperados. Existe un grupo de tratamiento y uno de control cuyos integrantes son elegidos aleatoriamente y sometidos a los mismos procedimientos y condiciones, con la salvedad de que uno de los grupos (el de tratamiento) está expuesto al programa a evaluar mientras que el otro no.	Las diferencias encontradas en ambos grupos se asocian con la intervención o el programa.
	Análisis costo-beneficio ¹⁴	Mide la eficiencia del programa, es decir, la relación entre los resultados obtenidos y los costos en los que se incurrió para la implementación del programa.	Provee información relevante acerca del uso eficiente de los recursos, por ejemplo, para la toma de decisiones presupuestaria.
	Monitoreo ¹⁵	Permite recolectar datos para el desarrollo de una evaluación. Consiste fundamentalmente en “examinar sistemáticamente un programa en términos de su cobertura y entrega de servicios” (Rossi, et al., 1979, p. 164).	Permite conocer con mayor detalle el funcionamiento del mismo; sin embargo, no permite observar el efecto o el beneficio derivado del programa sobre la población a la que va dirigido.
	Meta-análisis	Se utiliza cuando se cuenta con una serie de evaluaciones cuantitativas de un mismo programa/política cuyos resultados pueden combinarse con el fin de realizar inferencias más precisas acerca de los efectos del programa.	Requiere que las evaluaciones se hayan realizado con el rigor metodológico correspondiente para obtener información confiable.
Cuantitativo / Cualitativo	Evaluación de procesos	Se centra en la valoración de los distintos procedimientos de un programa. Su principal objetivo es identificar aquellas variables procedurales cuyo bien o mal funcionamiento afecta la obtención de resultados.	Se obtiene información acerca de distintas variables procedurales como por ejemplo: las características del personal que presta el servicio, la calidad de los servicios prestados, entre otras.
Cualitativo	Etnografía	Su objetivo principal es estudiar a fondo un grupo de personas en particular. El investigador que realiza la etnografía se involucra de manera directa en dicho grupo con el fin de captar y analizar las características del mismo.	La evidencia obtenida permite identificar características especiales de un grupo en relación con el objetivo de evaluación. Es recurrente para identificar cómo las características de un grupo en específico influyen en los resultados de un programa/política.
	Observación participante ¹⁶	Es un mecanismo para obtener información detallada de un programa a partir de la presencia de un actor con un rol relevante dentro del mismo.	Es posible obtener información muy específica acerca del programa; sin embargo, en la literatura se han discutido las implicaciones éticas de que exista un actor proveyendo información sin el consentimiento del resto de los involucrados.
	Estudio de caso	Se enfoca en el estudio de un elemento en particular con el propósito de realizar un análisis profundo.	Los hallazgos derivados del estudio de caso son válidos únicamente para éste, de ahí surge la principal crítica a esta herramienta, al señalar que los hallazgos no son generalizables pues las particularidades del caso influyen en los resultados obtenidos.

Fuente: Elaboración propia basada en Bennet (2003); Owen & Rogers (1999) Posavac & Carey (1989); Rossi & Freeman (1993); Stufflebeam (2001) y Weiss (1998a).

de cada una de ellas es concisa y, por consiguiente, limitada. Cada una de las herramientas presentadas ha sido desarrollada de manera individual por los especialistas en la materia por lo que existe una amplia literatura acerca de las mismas.

14 Ver: Levin, H. M. (1987). Cost-benefit and cost-effectiveness analyses. *New Directions for Program Evaluation* (34), 83-99. Levine, V. (1981). The role of outcomes in cost-benefit evaluation. *New Directions for Program Evaluation* (9), 21-40.

15 Ver: Cunill-Grau, N. & Ospina, S. M. (2012). Performance measurement and evaluation systems: Institutionalizing accountability for governmental results in Latin America. *New Directions for Evaluation* (134), 77-91

16 Ver: Horayangkura, V. (1989). Observer-as-participant method of data gathering. *New Directions for Program Evaluation*, (42), 65-74.

El nivel de sofisticación que han alcanzado muchos de los métodos de evaluación ha generado una discusión acerca de la predominancia de unos sobre otros:

«Parece ser que algunas agencias siguen apegándose a los métodos de siempre [...] esto puede ser un hábito (la familiaridad con un método respecto de otro), o a la falta de voluntad para tomar el riesgo de emparejar los métodos con mayor precisión a las preguntas de evaluación, o inclusive, como algunos sugieren, aun la incapacidad de lidiar con la complejidad [...], aún así en algunas agencias parece existir una mayor comprensión de las fortalezas y debilidades de ciertos métodos para responder una pregunta en particular y una mayor disposición para utilizar métodos mixtos que utilicen la fortaleza de uno o varios para compensar las deficiencias de otros» (Chelimsky, 2012, p. 79).

Esto puede responder a razones organizacionales como señala Chelimsky (2012) pero también a cuestiones políticas. El componente político de la evaluación está presente en todo el proceso y tiene consecuencias importantes¹⁷. Constriñéndose al ámbito metodológico, la influencia política del proceso puede incidir en que se prefiera un método sobre otro porque éste produce resultados más favorables políticamente (Bovens *et al.*, 2006) y, naturalmente, nadie quisiera ser señalado como responsable de la ineffectividad de un programa.

En este sentido, la selección metodológica genera «el dilema entre reconocer la realidad política de la evaluación y conservar el valor simbólico de la neutralidad» (Maynard-Moody *et al.*, 1990, p. 20). Los evaluadores tienen la obligación de emitir recomendaciones acerca del desempeño de un programa basándose en los hallazgos obtenidos y en el contexto político en el que están inmersos, «no realizar esta distinción entre esfuerzos exitosos y fallidos puede socavar a los programas sociales en lo general y a la evaluación en lo particular» (Eddy & Berry, 2009, p. 375). Mostrarse parcial frente a uno o varios de los implicados en la evaluación puede erosionar la credibilidad construida y constituir una amenaza a la independencia (Mohan, 2014).

El reto principal pareciera ser entonces impulsar que las instancias encargadas de la evaluación profundicen en el análisis de las metodologías existentes con el propósito de identificar aquella o aquellas que se ajustan mejor a las necesidades de la evaluación y a su contexto político. Esto incluye también reflexionar acerca de la naturaleza del objeto de evaluación, el propósito y audiencia de la misma, así como acerca de las herramientas metodológicas disponibles. El rol del evaluador es adoptar una posición de liderazgo para «comunicarse con las distintas audiencias, crear consenso, guiar el proceso de toma de decisiones y defender el diseño de la evaluación de críticas potenciales» (Braverman, 2012, p. 111).

17 Para mayor referencia acerca de la influencia política de la evaluación se sugiere ver: Bovens, M., t'Hart, P. & Kuipers, S. (2006). The politics of evaluation. En *The Oxford Handbook of Public Policy* (pp. 319–335). Oxford: Oxford University Press; Palumbo, D. (1987). *The politics of program evaluation*. California: Sage Publications y Weiss, C. H. (1993). Where politics and evaluation research meet. *American Journal of Evaluation*, 14 (1), 93–106.

6. Metodología y Evaluación: algunas reflexiones sobre el caso de México

La política de evaluación implementada por el Gobierno Federal permite establecer un vínculo más directo entre los aspectos teóricos y prácticos de la metodología de evaluación. Como se argumenta en este artículo, la selección metodológica es una decisión que trasciende las bondades técnicas y científicas que ofrece cada enfoque; aspectos como el *para qué, para quién y el qué evaluar* son centrales en la decisión.

Con base en el marco de referencia propuesto es posible identificar algunos aspectos relevantes de la política de evaluación mexicana en donde se observan distintos propósitos de la evaluación. De manera general, el Consejo Nacional de Evaluación de la Política de Desarrollo Social (CONEVAL) establece como propósito general de la evaluación: «mejorar constantemente su desempeño y conocer cuáles de las acciones son o no efectivas para resolver los grandes problemas sociales y económicos que todavía aquejan al País» (CONEVAL, 2012). Sin embargo, también es posible encontrar otros propósitos derivados de los instrumentos metodológicos desarrollados por esta organización, centrados en una lógica instrumental. Por ejemplo, la evaluación de diseño establece como propósito «evaluar el diseño del programa con la finalidad de proveer información que retroalimente su diseño, gestión y resultados» (CONEVAL, 2015, p. 2).¹⁸

Uno de los aspectos que llama la atención es que no existe un criterio explícito para la planeación de las evaluaciones. El Programa Anual de Evaluación (PAE), instrumento rector de este proceso, no hace evidente las razones por las que un grupo de programas es evaluado con un instrumento determinado y en un momento específico. Intuitivamente, existe una lógica vinculada con el ciclo de vida de un programa. Así, la evaluación de diseño, por ejemplo, está dirigida a programas nuevos o de reciente creación, mientras que a aquéllos que llevan un tiempo operando (también indefinido) se les mandata una evaluación de procesos. Fuera de esta deducción, existen pocos elementos para conocer con mayor detalle el proceso de planificación de las evaluaciones y cómo se vincula éste con la elección metodológica.

En términos de la audiencia, dado el propósito instrumental de la evaluación, se identifica que los usuarios principales son los operadores de los programas y, a un nivel más general, los tomadores de decisiones dentro y fuera de la Administración Pública Federal (APF). La Evaluación Específica de Desempeño (EED), por ejemplo, enfatiza que la información generada a través de este instrumento está dirigida a contribuir a la toma de decisiones (CONEVAL, 2012, p. 6). Aunque no se especifica quiénes son los tomadores de decisión a los que están dirigidos los hallazgos de la evaluación, la información derivada de ésta constituye un insumo importante para los legisladores en lo relativo al proceso presupuestario. Las características de esta información están adaptadas a esta audiencia. El reporte *Consideraciones para el Proceso Presupuestario*, por ejemplo, tiene una composición distinta a la de un reporte de evaluación. Provee datos puntuales acerca del desempeño de los programas a través de indicadores y de una valoración sintética, lo que permite a los legisladores tener una idea más clara de cuál ha sido el comportamiento de un programa durante un determinado ejercicio fiscal. Esto muestra

¹⁸ El propósito instrumental de la evaluación también se observa en la evaluación de procesos y específica de desempeño.

que existe una adaptación de los productos de evaluación en función de la audiencia; sin embargo, no se advierte ningún vínculo entre esta y la decisión metodológica.

Respecto del objeto de evaluación es posible identificar al menos dos niveles en el caso mexicano. El primero tiene como base los programas presupuestarios como objetos de evaluación; el segundo, mucho menos recurrente, es el que toma las políticas como objeto y que el CONEVAL ha denominado evaluaciones estratégicas.¹⁹ Esta es quizá una de las principales debilidades identificadas en la política de evaluación del Gobierno Federal al concentrar gran parte de los esfuerzos y recursos en el nivel programático:

«No existen las bases institucionales y normativas para llevar a cabo una evaluación sistémica integrada tanto horizontal como verticalmente, conciliando los niveles macro (políticas), meso (programas y dependencias), y micro (desempeño de los servidores públicos) [...] se corre el riesgo de disponer de información incompleta para la toma de decisiones, la asignación presupuestal y la rendición de cuentas» (González Arredondo, 2008, p. 241).

Un punto interesante acerca de México son los elementos importados del modelo chileno en el desarrollo del presupuesto basado en resultados el cual excluyó, por ejemplo, el componente del desempeño de funcionarios por lo que el diseño acotó el objeto de evaluación a programas e indicadores (González Gómez & Velasco, 2014). Esto es observable también en otros países de América Latina. En Colombia, Costa Rica y Chile gran parte del enfoque de evaluación está concentrado en el nivel programático.²⁰ En este punto tampoco se observa una vinculación clara con la decisión metodológica, programas de naturaleza muy distinta utilizan el mismo instrumento para la evaluación. Por ejemplo, el PAE 2011 estableció que el Programa Nacional de Lectura (PNL) realizara una Evaluación de Consistencia y Resultados (ECR), del mismo modo, se mandató la misma evaluación para el subsidio que se otorga a las instituciones estatales de cultura. Aunque ambos son programas presupuestarios operados por la Secretaría de Educación Pública (SEP), como objetos de evaluación presentan características distintas. Una diferencia fundamental entre ambos programas son las reglas de operación. Mientras que en el PNL estas reglas establecen claramente los objetivos y procesos de este programa, en el caso del subsidio a las instituciones estatales de cultura esto no es aplicable. Lo anterior lleva a reflexionar en qué medida los instrumentos actuales implementados para la evaluación se adaptan a las necesidades reales de los programas y generan insumos de utilidad para los operadores.

Este punto se vincula directamente con el último elemento del marco de referencia: la metodología. CONEVAL ha llevado a cabo una importante tarea en el desarrollo de instrumentos de evaluación y en el establecimiento de términos de referencia para la aplicación de evaluaciones estandarizadas como la EED y la ECR. Aun cuando esto facilita, como se señaló, la implementación de la política de evaluación es importante

19 Estas evaluaciones «diagnostican y analizan una problemática pública, así como la respuesta gubernamental para atenderla». Fuente: Consejo Nacional de Evaluación de la Política de Desarrollo Social. (2012). Recuperado de: <http://www.coneval.gob.mx/evaluacion/Paginas/Evaluacion.aspx>

20 Ver: Departamento Nacional de Planeación. (2010). *15 años del Sistema Nacional de Evaluación de Gestión y Resultados—Sinergia: Una mirada desde las evaluaciones de política pública más relevantes*; Ministerio de Planificación Nacional y Política Económica de Costa Rica. (2012). *Manual Gerencial para el Diseño y Ejecución de Evaluaciones Estratégicas de Gobierno*; y Olavarría Gambi, M. (2012). La evaluación de programas en Chile: Análisis de una muestra de programas públicos evaluados. *Revista del CLAD Reforma y Democracia*, 54.

reconocer que a su vez limita al evaluador a un esquema predeterminado que reduce la gama de métodos que puede aplicar. Muchos de estos instrumentos están basados en análisis documental, por lo que aspectos más específicos de un programa no son fácilmente observables. De nuevo, esto evidencia la desconexión entre la selección del método y variables como el propósito, objeto y audiencia de la evaluación.

A manera de conclusión y considerando las lecciones del caso mexicano, es posible decir que la elección de una metodología para evaluar políticas públicas tiene implicaciones mucho más amplias que la fortaleza técnica y el rigor científico de las herramientas. Si bien cada una de éstas ha sido desarrollada y discutida en la literatura, la evaluación se ve afectada por las decisiones que se toman durante todo el proceso de política pública y esto tendrá un efecto importante en los resultados obtenidos. Quienes tienen la facultad de comisionar e implementar las evaluaciones deben entenderlas como ejercicios complejos y multifactoriales. La evaluación no debe perder el sentido como actividad de valoración de un objeto, por el contrario, considerando los sujetos que la definen, participan y la modifican, debe constituir un esfuerzo de generación de información y herramientas útiles para la toma de decisiones, así como para la mejora del mismo proceso de evaluación.

7. Referencias Bibliográficas

- Bamberger, M., Rao, V. & Woolcock, M. (2010). Using mixed methods in monitoring and evaluation: Experiences from international development. Policy Research working paper ; no. WPS 5245. Washington, DC: World Bank.
- Bennett, J. (2003). *Evaluation methods in research*. Londres: Continuum.
- Berriet-Sollic, M., Labarthe, P. & Laurent, C. (2014). Goals of evaluation and types of evidence. *Evaluation*, 20 (2), 195-213.
- Boardman, C. (2014). Assessing Governance: The importance of evaluating policy outcomes in national mission areas. *Governance*, 27 (3), 519-526.
- Bogensneider, K. & Corbett, T. J. (2010). *Evidence-Based Policymaking: Insights from Policy-Minded Researchers and Research-Minded Policymakers*. Nueva York, Routledge.
- Bovens, M., t'Hart, P. & Kuipers, S. (2006). The politics of evaluation. En *The Oxford Handbook of Public Policy* (pp. 319-335). Oxford: Oxford University Press.
- Bovens, M. (2007). Analysing and Assessing Accountability: A Conceptual Framework. *European Law Journal*, 13 (4), 447-468.
- Braverman, M. T. (2012). Negotiating Measurement: Methodological and Interpersonal Considerations in the Choice and Interpretation of Instruments. *American Journal of Evaluation*, 34 (1), 99-114.
- Bryson, J. M., Crosby, B. C. & Bloomberg, L. (2014). Public Value Governance: Moving Beyond Traditional Public Administration and the New Public Management. *Public Administration Review* 74 (4), 445-456.

- Campbell, D. M. Redman, S., Jorm, L., Cooke, M. Zwi, A. B. & Rychetnik, L. (2009). Increasing the use of evidence in health policy and views of policymakers and researchers. *Australia and New Zealand Health Policy*, 6, 21.
- Cejudo, G. (2011). De las recomendaciones a las acciones: El uso de la evaluación en la mejora del desempeño del gobierno. En G. Cejudo & C. Maldonado (Eds.), *De las recomendaciones a las acciones: La experiencia del Premio 2011 Programas Federales Comprometidos con el Proceso de Evaluación* (pp. 11-28) México, D.F.: CIDE-SFP.
- Chelimsky, E. (1987). What Have We Learned about the Politics of Program Evaluation? *Educational Evaluation and Policy Analysis* 9(3), 199-213
- (2012). Valuing, evaluation methods, and the politicization of the evaluation process. *New Directions for Evaluation* (133), 77-83.
- Chen, H. T. (1994). Current Trends and Future Directions in Program Evaluation. *American Journal of Evaluation*, 15 (3), 229-238.
- Consejo Nacional de Evaluación de la Política de Desarrollo Social (CONEVAL) (2012). *Modelo de Términos de Referencia para la Evaluación Específica de Desempeño*. Recuperado de http://www.coneval.gob.mx/rw/resource/coneval/EVALUACIONES/NORMATIVIDAD/TdR_EED_2012/EED_TdR_2012_2013.pdf
- (2013). *Modelo de Términos de Referencia para la Evaluación de Procesos*. Recuperado de http://www.coneval.gob.mx/rw/resource/coneval/EVALUACIONES/NORMATIVIDAD/TdR_Procesos_2013/TdR_Procesos%202013_Oficio.pdf
- (2015). *Modelo de Términos de Referencia para la Evaluación Específica de Diseño*. Recuperado de http://www.coneval.gob.mx/rw/resource/coneval/eval_mon/normatividad_matriz/Modelo_de_terminos_de_referencia_evaluacion_diseno_final.pdf
- Consejo de Evaluación-Universidad de Chile. (2011). *Modelo de Diseño y Evaluación de Procesos Estratégicos*. Santiago de Chile.
- Contandriopoulos, D. & Brousselle, A. (2012). Evaluation models and evaluation use. *Evaluation* 18 (1), 61-77.
- Cousins, J. B., Goh, S. C., Elliott, C. J. & Bourgeois, I. (2014). Framing the Capacity to Do and Use Evaluation. *New Directions for Evaluation* (141), 7-23.
- Creswell, J. W. (1994). *Research Design: Qualitative and Quantitative Approaches*. Londres: Sage Publications.
- Cronin, P., Ryan, F. & Coughlan, M. (2008). Undertaking a literature review: A step-by-step approach. *British Journal of Nursing* 17 (1), 38-43.
- Cunill-Grau, N. & Ospina, S. M. (2012). Performance measurement and evaluation systems: Institutionalizing accountability for governmental results in Latin America. *New Directions for Evaluation* (134), 77-91.

- Davidson, E. J. (2005). *Evaluation Methodology Basics: The Nuts and Bolts of Sound Evaluation*. Londres: Sage Publications.
- Departamento Nacional de Planeación (2010). *15 años del Sistema de Evaluación de Gestión y Resultados-Sinergia: Una mirada desde las evaluaciones de políticas más relevantes*. Recuperado de <http://siare.clad.org/siare/innotend/evaluacion/colombia/c1.pdf>
- Durand, R., Decker, P. J. & Kirkman, D. M. (2014). Evaluation Methodologies for Estimating the Likelihood of Program Implementation Failure. *American Journal of Evaluation*, 35 (3), 404-418.
- Eddy, R. M. & Berry, T. (2009). The Evaluator's Role in Recommending Program Closure. *American Journal of Evaluation*, 30 (3), 363-376.
- Feinstein, O. N. (2002) Use of Evaluation and the Evaluation of their Use. *Evaluation*, 8 (4), 433-439.
- Gaarder, M. M. & Briceño, B. (2010). *Institutionalisation of Government Evaluation: Balancing Trade-Offs. International Initiative for Impact Evaluation (3ie)*.
- Gertler, P. J. Martínez, S. Premand, P. Rawlings, L. B. & Vermeesch, C, M. J. (2011). *La evaluación de impacto en la práctica*. Washington, D. C.: Banco Mundial.
- González Arredondo, A. (Ed.) (2008). *¿Gobernar por resultados? Implicaciones de la política de evaluación del desempeño del gobierno mexicano*. México, D. F.: GESOC, A. C.
- González Gómez, J. & Velasco Sánchez, E. (2014). La evolución del presupuesto basado en resultados en México, 2004-2012: transferencia e implantación de una política pública. *Revista del CLAD Reforma y Democracia* (58).
- Greene, J. C., Lipsey, M. W., Schwandt, T. A. Smith, N. L. & Tharp, R. G. (2007). Method choice: Five discussant commentaries. *New Directions for Evaluation* (113), 111-127.
- Gris-Legorreta, P. C. (2011). La evaluación de políticas como instrumento para la rendición de cuentas. *Trimestre Fiscal*, 99, 89-103.
- (2013) Escapar del olvido. La promoción del uso de la evaluación en México. *Revista Este País*, 271.
- Hanberger, A. (2006). Evaluation of and for Democracy. *Evaluation*, 12 (1), 17-37.
- Hansen, H. (2005) Choosing Evaluation Models A Discussion on Evaluation Design. *Evaluation*, 11 (4), 447-462.
- Henry, G. T. & Mark, M. M. (2003). Beyond Use: Understanding Evaluation's Influence on Attitudes and Actions. *American Journal of Evaluation*, 24 (3), 293-314.
- Hertting, N. & Vedung, E. (2012). Purposes and criteria in network governance evaluation: How far does standard evaluation vocabulary takes us? *Evaluation*, 18 (1), 27-46.
- Horayangkura, V. (1989). Observer-as-participant method of data gathering. *New Directions for Program Evaluation*, 42, 65-74.

- Huberman, A. M. & Miles, M. B. (2002). *The qualitative researcher's companion*. Londres: Sage Publications.
- Johnson, R. B. & Onwuegbuzie, A. J. (2004). Mixed Methods Research: A Research Paradigm Whose Time has Come. *Educational Researcher*, 33 (7), 14-26.
- Jung, C. S. (2014). Why are goals important in the public sector: Exploring the benefits of Goal Clarity for Reducing Turnover Intention. *Journal of Public Administration Research and Theory*, 24 (1), 209-234.
- Kromann, J., Groszyk, W. & Bühler, B. (2001) Outcome-focused Management and Budgeting. *OECD Journal on Budgeting*, 1 (4), 7-34.
- Kusek, J. Z. & Rist, R. C. (2004). *Ten steps to a results-based monitoring and evaluation system: A handbook for development practitioners*. Washington, D. C.: Banco Mundial.
- Lawrenz, F., Gullickson, A., & Toal, S. (2007). Dissemination. *American Journal of Evaluation*, 28 (3), 275-289.
- Levin, H. M. (1987). Cost-benefit and cost-effectiveness analyses. *New Directions for Program Evaluation* (34), 83-99.
- Levine, V. (1981). The role of outcomes in cost-benefit evaluation. *New Directions for Program Evaluation* (9), 21-40.
- Leviton, L. C. & Hughes, E. F. X. (1981). Research on the utilization of evaluations: A review and synthesis. *Evaluation Review*, 5 (4), 525-548.
- Lundin, M. & Öberg, P. (2014). Expert knowledge use and deliberation in local policymaking. *Policy Sciences*, 47 (1), 25-49.
- Majone, G. (2005). *Evidencia, argumentación y persuasión en la formulación de políticas públicas*. México, D. F.: Fondo de Cultura Económica.
- Mason, J. (2005). *Qualitative Researching*. Londres: Sage Publications.
- Matthews, B. & Ross, L. (2010). *Research Methods: A practical guide for the social sciences*. Harlow: Longman.
- Maynard-Moody, S., Musheno, M. & Palumbo, D. (1990). Street-Wise Social Policy: Resolving the Dilemma of Street-Level Influence and Successful Implementation. *Political Research Quarterly*, 43 (4), 833-848.
- Ministerio de Planificación Nacional y Política Económica de Costa Rica. (2012). *Manual Gerencial para el Diseño y Ejecución de Evaluaciones Estratégicas de Gobierno*. Recuperado de https://documentos.mideplan.go.cr/alfresco/d/d/workspace/SpacesStore/9d96f810-3b89-4fd6-8701-d36327f02cc1/Manual_Gerencial_para_Dise%C3%B1o_Ejecucion_Evaluaciones_Estrategicas_Gobierno.pdf?guest=true
- Mohan, R. (2014). Evaluator Advocacy: It is all in a day's work. *American Journal of Evaluation*, 35 (3), 397-403.

- Neuman, A., Shahor, N., Shina, I. Sarid, A. & Saar, Z. (2013). Evaluation utilization research-Developing a theory and putting it to use. *Evaluation and Program Planning*, 36 (1), 64-70.
- Newcomer, K. E. (1997). Using performance measurement to improve programs. *New Directions for Evaluation* (75), 5-14.
- Olavarría Gambi, M. (2012). La evaluación de programas en Chile: Análisis de una muestra de programas públicos evaluados. *Revista del CLAD Reforma y Democracia*, 54.
- Owen, J. M. & Rogers, P. (1999) *Program Evaluation. Forms and Approaches*. Melbourne: Sage Publications.
- Palumbo, D. (1987). *The politics of program evaluation*. California: Sage Publications.
- Patton, M. Q. (1986). *Utilization-focused evaluation*. Londres: Sage Publications.
- (2002). *Qualitative Research & Evaluation Methods*. Londres: Sage Publications.
- Picciotto, R. (2005). "Policy Coherence and Development Evaluation: Issues and Possible Approaches." En *Fostering Development in a Global Economy: A Whole of Government Perspective*. OECD.
- Posavac, E. J. & Carey, R. G. (2003). *Program evaluation: Methods and case studies*. Nueva Jersey: Prentice Hall.
- Ramdhani, A., Ramdhani, M. A. & Syakur Amin, A. (2014). Writing a Literature Review Research Paper: A step-by-step approach. *International Journal of Basic and Applied Science*, 3 (1), 47-56.
- Rao, V. & Woolcock, M. (2003). Integrating Qualitative and Quantitative Approaches in Program Evaluation. En F. Bourguignon & L. A. P. Da Silva (Eds.), *The impact of economic policies on poverty and income distribution: Evaluation techniques and tools*. Washington, D. C.: Banco Mundial.
- Rhodes, R. A. W. (1997). *Understanding governance: Policy networks, governance, reflexivity and accountability*. Buckingham: Open University Press.
- Rossi, P. H., Freeman, H. E. & Rosenbaum, S. (1979). *Evaluation: A systematic approach*. Londres: Sage Publications.
- Sanderson, I. (2001). Performance Management, Evaluation and Learning in 'Modern' Local Government. *Public Administration*, 79 (2), 297-313.
- Santiso, C. (2015). Governing to Deliver: Three Keys for Reinventing Government in Latin American and the Caribbean. *Governance*, 28 (2), 123-126.
- Secretaría de Salud. (2013). *Protección contra Riesgos Sanitarios. Evaluación de Procesos*. México, D. F. Recuperado de http://www.cofepris.gob.mx/Transparencia/Documents/INFORME_EVALUACION_FIN.pdf

- Shadish, W.R. & Rindskopf, D. M. (2007). Methods for evidence-based practice: Quantitative synthesis of single-subject designs. *New Directions for Evaluation* (113), 95-109.
- Stufflebeam, D. 2001. Evaluation Models. *New Directions for Evaluation* (89), 7-98.
- Talbot, C. (2010). *Performance in Government. The Evolving System of Performance and Evaluation Measurement, Monitoring, and Management in the United Kingdom*. Washington D.C., US: The World Bank.
- Uvalle, R. (2011). Las políticas públicas en el arquetipo de la gobernanza democrática. *Revista del CLAD Reforma y Democracia* (50).
- Vanlandingham, G. R. (2011). Escaping the Dusty Shelf: Legislative Evaluation Office' Efforts to Promote Utilization. *American Journal of Evaluation*, 32 (1), 85-97.
- Vartiainen, P. (2002). On the principles of comparative evaluation. *Evaluation*, 8 (3), 359-371.
- Watson, P. G. (2014). Can politicians afford knowledge? Evidence as affordance in new governance oversight arrangements. *Administration & Society*.
- Weible, C., Heikkila, T., deLeon, P. & Sabatier, P. (2012). Understanding and Influencing the Policy Process. *Policy Sciences*, 45 (1), 1-21.
- Weiss, C. H. (1993). Where politics and evaluation research meet. *American Journal of Evaluation*, 14 (1), 93-106.
- (1998a). *Evaluation: Methods for studying programs and policies*. Londres: Prentice Hall.
- (1998b). Have we learned anything new about the use of evaluation? *American Journal of Evaluation*, 19 (1), 21-33.
- Wholey, J. S. (2001). Managing for Results: Roles for Evaluators in a New Management Era. *American Journal of Evaluation*, 22 (3), 343-347.
- Widmer, T. & Neuenschwander, P. (2004). Embedding Evaluation in the Swiss Federal Administration. *Evaluation*, 10 (4), 388-409.
- Zapico-Goñi, E. (2007). Matching Public Management, Accountability and Evaluation in Uncertain Contexts. *Evaluation*, 13 (4), 421-438.
- Øyen, E. (1990). *Comparative Methodology: Theory and Practice in International Social Research*. Londres, Sage Publications.

8. Nota Metodológica

Este artículo adopta un enfoque cualitativo (Huberman & Miles, 2002; Mason, 2002) basado fundamentalmente en una revisión literaria y análisis de fuentes secundarias acerca de la metodología de evaluación. El desarrollo de un marco de referencia para vincular la metodología con el uso de la evaluación requiere un estudio detallado de la

literatura en esta materia. Una revisión literaria es «una síntesis objetiva y un análisis crítico de la literatura disponible y relevante acerca del objeto de estudio» (Cronin, Ryan & Coughlan, 2008, p. 38). Una de las finalidades de este instrumento es contribuir a la identificación de vacíos existentes en un área de conocimiento (Ramdhani, Ramdhani & Syakur Amin, 2014). Las variables que intervienen en la selección de la metodología permiten responder distintas interrogantes acerca de las políticas públicas (Davidson, 2005). Con base en el carácter cualitativo del artículo, la metodología se centra en la revisión, sistematización y análisis de la literatura más relevante acerca de la metodología para la evaluación. Este artículo sintetiza las principales discusiones acerca de la metodología y evaluación con el propósito de identificar aquellos aspectos considerados de mayor trascendencia para la selección metodológica. De manera específica, a continuación se describe el proceso:

1) Se realizó una revisión en revistas destacadas del campo de la evaluación (e.g. *Evaluation*, *New Directions for Evaluation*²¹) para identificar artículos que refirieran al aspecto metodológico de la evaluación. Se utilizaron como palabras de búsqueda “*methodology*” y “*methods*”. 2) Se seleccionaron aquellos artículos cuyo tema central estuviera directamente vinculado con la metodología de evaluación, priorizando aquellos más recientes. 3) Se identificaron elementos relacionados con la decisión metodológica para el desarrollo del marco conceptual que se propone en el presente artículo. La identificación de estas variables constituyó el elemento central para el desarrollo de la discusión acerca de la metodología de evaluación. 4) A partir del análisis de fuentes secundarias (informes de evaluación, documentos gubernamentales) se estableció el vínculo teórico-práctico de la discusión con el caso mexicano.

21 Estas publicaciones son consideradas referencias obligadas en el campo de la evaluación y están incluidas en los principales índices de publicaciones científicas. *Evaluation*, por ejemplo, presenta los siguientes datos en relación con su posición en estos índices: Ranking: 2013 SJR (SCImago Journal Rank) Score: 0.505 | 63/194 Development | 283/902 Sociology and Political Science (Scopus)