

# Usos, aportes y desafíos de la evaluación como instrumento de políticas públicas

## *Uses, contributions, and challenges of evaluation as a public policy instrument*

José Del Tronco Paganelli\*

Recibido: 07/04/2023 - Aceptado: 27/04/2023

### Resumen

El objetivo fundamental de este artículo es reflexionar y debatir, con base en la literatura especializada, los usos, aportes y desafíos de la evaluación en el ámbito de las políticas públicas. Partiendo de reconocer la pertinencia de esta actividad en los distintos espacios de la vida pública e institucional de una sociedad, en este documento se hace hincapié en aquellas nociones que toda persona vinculada al estudio o la práctica de las políticas públicas debe conocer. Asimismo, se analizan las implicaciones de considerar a la evaluación un instrumento de política pública que es relevante en las diversas fases del ciclo. Finalmente, se pregunta y analiza cómo se puede conciliar desde la evaluación las diversas racionalidades presentes en el proceso de toma de decisiones.

*Palabras clave:* aprendizaje, deliberación, evaluación, políticas públicas, racionalidad.

### Abstract

The aim of this document is to discuss the uses, contributions, and challenges faced by public policy evaluation. Acknowledging the importance of this activity in different parts of political and institutional life, the document focuses on concepts that any person related to the study or practice of public policy should know. The article also discusses the effects of considering evaluation as a policy instrument that is ubiquitous in different phases of policy process. Finally, it analyzes the role of public deliberation for reconciling diverse forms of rationality that are present in decision-making processes.

*Keywords:* deliberation, evaluation, learning, public policy, rationality.

---

\* Profesor-investigador de tiempo completo en FLACSO México. [jdetronco@flacso.edu.mx](mailto:jdetronco@flacso.edu.mx)

## Introducción: Ideas fuerza y propuestas para el debate en la evaluación de políticas

Este artículo parte de tres ideas centrales. La primera es que la evaluación es –o al menos debería ser– una actividad transversal al ciclo de políticas públicas. Desde la definición y diagnóstico del problema público hasta la ponderación del impacto del programa sobre la población beneficiaria, tal actividad ocupa un rol fundamental a lo largo del proceso de la política. La segunda refiere a que la evaluación de políticas públicas está atravesada de forma simultánea por los criterios –en ocasiones contradictorios– de racionalidad instrumental y factibilidad política. La tercera alude a que la evaluación resulta mucho más efectiva cuando se concibe como un instrumento que promueve el aprendizaje (*policy learning*) en lugar del enfoque punitivo que la considera exclusivamente como una estrategia de control de gestión y rendición de cuentas; en el primer caso, aumenta la responsabilidad y el compromiso con la mejora continua de las políticas, mientras que en el segundo (emparentado con la sanción) genera temor, obstaculiza la rendición de cuentas y, en el extremo, promueve la parálisis del proceso decisorio.

En el ámbito de las políticas públicas, la evaluación debe lidiar con dos dimensiones de una misma realidad: a) la valorativa, que refiere a los dilemas y objetivos políticos que todo programa se propone alcanzar, y b) la instrumental, que destaca el carácter técnico del proceso decisorio; la política seleccionada debe ser aquella capaz de instrumentar los medios idóneos -costo-efectivos- para maximizar el logro de los objetivos sociales planteados en el diseño (Fontaine 2015)–. Dicho de otra forma, la política implementada debe ser aquella que maximice el logro de objetivos y, a su vez, genere el menor rechazo social para fortalecer la legitimidad del gobierno frente a la ciudadanía (Subirats 2015).

Por lo general, ambas motivaciones son difíciles de alcanzar. Más aún, suelen ser parte de dos modos (alternativos en el extremo) de concebir el proceso de políticas. Aquellos resultados que postulan las comunidades académicas en sus “torres de marfil” (sin otras limitaciones que las del razonamiento científico en un mundo de factores controlados por la cláusula “*ceteris paribus*”), difieren tanto en términos semánticos como prácticos de las realidades cotidianas de los “burócratas”, sujetos a las restricciones de tiempo, recursos, jerarquías políticas y ciclos electorales del mundo real (Löfgren y Bickerton 2013, 132).

El escenario de la evaluación, por tanto, se define por la interacción entre los actores, las condiciones “objetivas” y los intereses involucrados. Así, la evaluación de políticas debe navegar con “bandera de corso” entre las aguas turbulentas de la deliberación política y las no mucho menos ambiciosas (y pretensivas) de la eficiencia y la efectividad.

Dada esta dualidad, en el artículo se discute cuáles son los usos, los abusos y la utilidad de la evaluación de políticas públicas, desde ambas perspectivas: la política y la burocrática; la vinculada con intereses e ideas de lo adecuado, y la emparentada con criterios racionales y evidencia científica. Expresado en otros términos, si bien los actores tratan de cumplir con las obligaciones involucradas en el rol institucional que desempeñan, las expectativas sobre lo adecuado suponen tanto una ética profesional como una identidad política (March y Olsen 2006). ¿Es posible compatibilizar ambas necesidades (la política y la de políticas)? En tal sentido, ¿puede la evaluación concebirse como un instrumento de aprendizaje, para la mejora y construcción de conocimiento en apoyo de un proceso decisorio con claros sesgos ideológicos? O, por el contrario, ¿es más rentable entenderla como una actividad punitiva, orientada al control y la rendición de cuentas? En términos políticos, ¿es deseable equivocarse y aprender, o es mejor prevenir los fallos de gobierno, a riesgo de contribuir a las parálisis administrativas por temor a la sanción jurídica o política?

Para problematizar estos interrogantes, se ha organizado el artículo en las siguientes secciones. En la primera se presenta la definición conceptual de evaluación y las particularidades de la evaluación de políticas. Seguidamente, se repasan los orígenes del *policy evaluation* para entender algunas cuestiones relacionadas con el estado actual de la disciplina. En el tercer apartado, se analiza la evaluación no solo como fase terminal del proceso decisorio, sino como instrumento de políticas públicas: qué tipo de instrumento es y qué aporta en términos de aprendizaje. Luego se plantea cuán importante es pensar (y utilizar) la evaluación en cuanto actividad transversal al proceso de políticas públicas: ¿qué lógica y aplicaciones tiene en cada etapa y cuáles son las preguntas detonantes para su aplicación? En la quinta sección, se analiza la importancia de los criterios de evaluación y se discute si es posible conciliar –a través de las prácticas deliberativas– las motivaciones políticas del proceso decisorio con los criterios técnicos que deben guiar el diseño –y, por tanto, la evaluación– de las políticas. En la parte final, se comparten algunos mitos y realidades de la evaluación de políticas, y algunas conclusiones que se espera aumenten el interés de lectoras y lectores en esta materia.

## ¿Qué es y para qué sirve la evaluación de políticas públicas?

En términos generales, evaluar es juzgar; ejercer algún tipo de valoración o ponderación sobre algún fenómeno (sociopolítico, en este caso) de acuerdo con determinados parámetros o criterios (Subirats y Gomà 1989). En el ámbito de las políticas públicas, hacer una evaluación es generar información de manera sistemática acerca

de las actividades, metas o resultados –presentes o futuros– de los problemas o programas en cuestión<sup>1</sup> (Patton 2002) para apoyar la toma de decisiones.

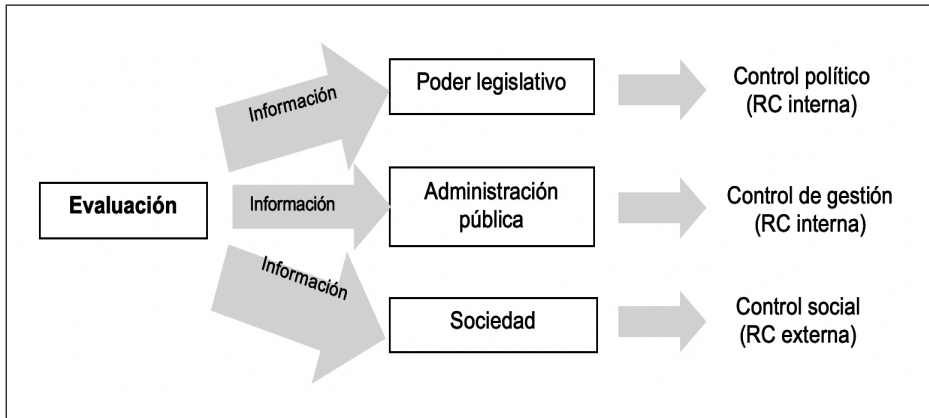
Las actividades de evaluación de programas públicos alcanzaron popularidad y difusión durante las décadas de los sesenta y los setenta en los Estados Unidos. Si bien ya existían antecedentes de esfuerzos gubernamentales para conocer los resultados de sus políticas –Page (2006, 207-208) ubica las primeras evaluaciones de políticas durante el gobierno del canciller Von Bismarck, en la Alemania de fines de siglo XIX– fue durante las administraciones de John F. Kennedy (1960-1963) y Lyndon Johnson (1963-1968) en Estados Unidos, que esta actividad alcanzó mayor reputación. El surgimiento de programas sociales destinados a acabar con el desempleo, el crimen y el deterioro urbano, así como la necesidad de generalizar el acceso a servicios médicos y educativos acarrearón una enorme cantidad de recursos humanos y financieros, por lo cual resultaba imperativo conocer sus resultados (Pressman y Wildawsky 1998).

Tres fueron las razones principales por las cuales era (es) necesario evaluar las políticas implementadas. En primer lugar, el gobierno debía informar al Congreso acerca de los resultados concretos de las leyes que allí se habían votado; dicho de otro modo, era y sigue siendo obligatorio para quienes gobiernan rendir cuentas de sus decisiones. En segundo término, era (es) imprescindible para los administradores públicos conocer cómo operan sus programas, los resultados que alcanzan y los recursos que insumen para el logro de sus objetivos (Subirats y Gomà, 1989). Finalmente, el tercer punto: en contextos democráticos, la opinión pública –acicateada por la actividad de los medios de comunicación– demanda transparencia en los procedimientos de gobierno tanto como eficiencia y efectividad en la utilización de los recursos públicos (Morales, Meneses y Alonso 2006). Bajo esta perspectiva, la evaluación es un instrumento orientado a “controlar” el carácter democrático y la efectividad del proceso decisorio.

---

1 En algunos casos, la evaluación sirve tanto para analizar los alcances de los programas públicos como para conocer el desempeño de las personas encargadas de operar dichos programas. En particular, los enfoques que centran su atención en la fase de implementación, para entender los resultados no deseados de las políticas públicas, parten de esta hipótesis: el desempeño de los recursos humanos constituye el factor clave para el éxito o el fracaso de las intervenciones públicas.

Figura 1. La evaluación y las dimensiones del control y la rendición de cuentas (RC)



Elaborado por el autor.

*Nota:* De acuerdo con Schedler (2008), la rendición de cuentas está compuesta por tres dimensiones: la información, la argumentación y la sanción. En tal sentido, la evaluación de políticas resulta una herramienta fundamental puesto que arroja información válida sobre los resultados de las intervenciones públicas, y permite una mayor y mejor argumentación cuando estos no son los esperados. Asimismo, genera evidencia que aumenta o disminuye la posibilidad de sanción (al menos cuando esta última se considera en su dimensión electoral).

En la actualidad, se admite abiertamente que la evaluación de programas públicos sirve tanto para un mejor conocimiento de los mismos por parte de la opinión pública como de los funcionarios de la administración, quienes son conscientes de la importancia de la información generada por la actividad evaluadora (Subirats y Gomà, 1989). Sin embargo, vale la pena preguntarse si es suficiente. Por un lado, los meta-objetivos de la evaluación suelen ser imprecisos. Sabemos qué tipo de preguntas responder para evaluar, pero no por qué –con qué motivación– realizamos la evaluación. ¿Evaluamos para sancionar a quien no se apega a cierto estándar? ¿Evaluamos para motivar y premiar ciertas conductas? ¿Evaluamos para aprender de los “fallos” del proceso de políticas?, ¿o simplemente lo hacemos para cumplir con los requisitos normativos de una nueva gestión pública, que quiere apearse –en las normas, más que en los hechos– a los estándares sugeridos por los organismos internacionales?

### *La evaluación como proceso culturalmente enraizado*

De acuerdo con la Comisión Europea, evaluar es “realizar juicios de valor sobre las intervenciones públicas de acuerdo con los resultados, impactos y necesidades que aquellas buscan satisfacer. Es una herramienta sistemática que provee evidencia rigurosa para sostener el proceso de toma de decisiones” (Lázaro 2015, 11).

En tal sentido, la evaluación suele ser vista como la fase terminal del proceso de toma de decisiones. La herramienta que genera información sobre los logros alcanzados por un programa, a partir del diseño de ciertos procedimientos, de la implementación de ciertos procesos y actividades, y de la ejecución de una cantidad determinada de recursos. Sin embargo, la evaluación es más que ello; es en sí misma un instrumento de política pública. Ello significa que la evaluación forma parte del conjunto de técnicas por medio de las cuales un gobierno se propone modificar –más directa o indirectamente– los comportamientos de aquellos –funcionarios o ciudadanos– a los que identifica como causantes del problema a resolver por la política (Bemelmans-Videc 1998; Knoepfel et al. 2007). Evaluar, en estos casos, es valorar en qué medida tales cambios fueron generados exclusivamente por las acciones de política pública, y no por otras condiciones (OECD 2020).

Por lo general, los gobiernos llevan adelante evaluaciones para cumplir con un amplio rango de objetivos.<sup>2</sup> Por tanto, cuando pensamos en los usos de la evaluación, no podemos perder de vista que es un proceso culturalmente enraizado. Dicho de otro modo, la forma de concebir, aplicar y entender la evaluación está influenciada por las normas y los valores que guían las percepciones sociales y sentidos comunes de lo que es importante, deseable y valioso en el contexto donde la evaluación tiene lugar. Actuar de forma “apropiada” significa hacerlo de acuerdo con las prácticas institucionalizadas de una colectividad, que comparte significados sobre lo que es verdadero, razonable, correcto o valioso para esa comunidad, en ese lugar y en ese momento de la historia (March y Olsen 2006, 690).

Es aquí donde las ideas de lo adecuado adquieren importancia como criterios o estándares, que devienen bases para la toma de decisiones, y se evalúan los alcances y resultados de las mismas. De acuerdo con Peter Hall (1993), los encuadres (*framing*) de ideas y criterios dominantes determinan no sólo los objetivos e instrumentos, sino la misma naturaleza de los problemas a ser resueltos según estas nuevas miradas. De tal manera, los nuevos acuerdos paradigmáticos establecen tanto los (nuevos) fines sociales a alcanzar, así como los procedimientos legales y procesos de toma de decisiones a seguir para su obtención.

Dado que la identificación de una situación problemática tanto como la designación de públicos objetivo (causantes del y perjudicados por el problema público) son el resultado de una construcción sociopolítica (Jerit 2008), las ideas y los encua-

---

2 Medir los resultados alcanzados y los recursos insumidos para obtenerlos; promover evidencia sobre el funcionamiento del proceso de toma de decisiones en sus diversas fases o etapas; medir el impacto de las acciones gubernamentales en la mejora de la calidad de los servicios brindados, la satisfacción de los usuarios o beneficiarios, o la disminución del problema público en cuestión; aumentar la confianza y el apoyo de la ciudadanía gracias a una mayor transparencia en la publicitación de los resultados del gobierno, o reformular políticas públicas cuando sus resultados no hayan sido los esperados (OECD 2020, 6).

dres discursivos son herramientas valiosas para derribar los acuerdos previamente existentes y construir consensos en torno a valores y objetivos de las políticas públicas. Como consecuencia de ello, la evaluación (ejecutada tanto para diagnosticar los problemas como para valorar el impacto de las políticas) también lo es.

## Los usos y utilidades de la evaluación como instrumento de políticas

Como se apuntó previamente, la evaluación es más que una etapa del ciclo de políticas; es fundamentalmente un instrumento de la política pública. “Un instrumento de política pública es un dispositivo que es tanto técnico como social, que organiza relaciones sociales específicas entre el estado y aquellos a los que está dirigido, según las representaciones y los significados que conlleva” (Lascoumes y Le Galés 2019, 216).

El concepto de “instrumentos de política pública” supone que todo gobierno necesita de estos dispositivos tecno-políticos para incidir en la realidad social. Sin embargo, hasta hace algunas décadas, estos eran concebidos como componentes relativamente estáticos y exclusivos de los gobiernos que eran utilizados como recursos o capacidades necesarios para dar respuesta a diversos asuntos problemáticos incorporados a la agenda (Hood 1986).<sup>3</sup> En la actualidad, el concepto de “instrumentos” se emparenta con el proceso de diseño de acciones públicas (en las que intervienen actores diversos) por medio de las cuales se intenta modificar comportamientos sociales definidos como causantes de un problema social (Knoepfel et al. 2007; Howlett 2011).

Considerar la evaluación como un instrumento de las políticas públicas supone entenderla como una actividad –a la vez técnica y social– orientada a la mejora de los procesos gubernamentales de provisión de bienes y servicios públicos. Esta mejora, se supone, será resultado de la producción y aplicación de información sistemática (generada por las diversas formas y metodologías de evaluación) acerca de los problemas públicos que el gobierno decide atender, así como de las respuestas gubernamentales puestas en práctica para su atención. Dicho de otra forma, cuanto más sabemos de aquello que queremos resolver (sus causas, efectos y representaciones sociales), y de la forma en la que lo estamos haciendo (el grado de cambio que las acciones del gobierno están generando en el problema, así como sus efectos secundarios no deseados), dispondremos de más evidencia -válida y confiable- sobre la (in)efectividad de las políticas y la (in)satisfacción de aquellos a quienes van dirigidas (OECD 2020).

---

3 Autoridad, modalidad, tesoro y organización son las cuatro herramientas de gobierno identificadas por Christopher Hood en su trabajo seminal de la década de los ochenta.

En términos políticos, los instrumentos no son herramientas neutrales, sino portadores de valores e ideas, pero en términos técnicos, la idoneidad de los instrumentos dependerá de su capacidad para transformar las condiciones que propician los problemas públicos. Por ello, en los estudios contemporáneos de política pública, el diseño es la etapa asociada a la construcción y calibración de los instrumentos, y al establecimiento de estándares para su ejecución.

De acuerdo con Howlett, la actividad de diseño de políticas supone un componente sustantivo y uno procedimental. El primero se relaciona con la teoría de cambio que da sustento a la elección de los instrumentos para la solución del problema público, mientras que el componente procedimental se refiere a las condiciones, procesos y actividades que hacen posible la coordinación de los actores que participan de la implementación (Howlett 2011). Dicho en otros términos, de nada sirve diseñar instrumentos capaces de resolver problemas, si no están presentes las condiciones de factibilidad (sociales, políticas, financieras y técnicas) que hagan posible su ejecución. En tal sentido, la evaluación de políticas públicas puede concebirse como un “superinstrumento” de política pública. Por un lado, forma parte de los instrumentos procedimentales que delimitan las condiciones en que una política pública debe ser formulada y ejecutada; por el otro, genera información relevante y necesaria para valorar la idoneidad de los instrumentos sustantivos, en función de sus resultados.

Los instrumentos “sustantivos” responden a una hipótesis de intervención. Es decir, se parte de un diagnóstico causal de los elementos que propician, promueven o producen una situación problemática. Una vez constatada dicha relación, se postulan una serie de acciones cuyo efecto esperado es mitigar, neutralizar o eliminar los comportamientos generadores del problema público, de forma de lograr su mitigación, neutralización o eliminación.

Una revisión de la literatura en la que se analizan los instrumentos sustantivos a partir de las lógicas de diseño, nos permite identificar –al menos– cuatro tipos de instrumentos:

- 1) los instrumentos punitivos, o de tipo “garrote”: aquellos que buscan cambiar comportamientos por medio de la proscripción (prohibición) de comportamientos. Distinguen lo permitido de lo prohibido, a través de prescripciones o proscripciones de comportamientos. Las leyes o reglamentaciones de diverso alcance son las formas típicas de regulación;
- 2) los instrumentos remunerativos o de tipo “zanahoria”: aquellos que –sin establecer prohibiciones– aumentan o disminuyen el costo de ciertas conductas, para quienes deciden ejercerlas. Los impuestos, los subsidios o las becas son ejemplos de este tipo de instrumentos;



- 3) los instrumentos informativos o de tipo “sermón”: brindan recomendaciones a la población objetivo para que esta pueda calcular de mejor manera las consecuencias de sus actos. Dicho de otra forma, se promueve la racionalidad de los actores sin mediar condicionamientos externos a la conducta (prohibiciones o remuneraciones) en un marco de mayor libertad. Prescriben, como las recetas médicas, las conductas más acordes con la salud del cuerpo social. Ejemplos de este tipo de instrumentos son las campañas orientadas a demostrar el beneficio de ciertos comportamientos como la separación de residuos, o el uso de condón, cuando se quieren propiciar resultados sociales considerados benéficos (disminuir el costo de los procesos de desecho de residuos) o evitar situaciones indeseables (reducir la tasa de embarazos en edades tempranas).
- 4) los instrumentos de tipo *nudge*: popularizados en las últimas dos décadas, este tipo de instrumentos adquieren una lógica similar a los informativos, pero no apelan a la racionalidad de los individuos, sino a su sistema “automático” de toma de decisiones. Se caracterizan por modificar intencionalmente la arquitectura de las decisiones de aquellos cuyos comportamientos se quieren orientar, generando opciones por defecto (*by default*) en las que la decisión socialmente indeseable (de acuerdo con la perspectiva política de quien diseña el instrumento) es la más costosa de tomar (Sunstein y Thaler 2018).

Lo interesante de esta clasificación es que los instrumentos de tipo “zanahoria” o “garrote” requieren de un incentivo externo para orientar la conducta de acuerdo con los objetivos del gobierno. Por su parte, los “sermones” aumentan la libertad de las personas y apuntan a que su conducta se vea informada por motivaciones internas y no por incentivos externos (“separo la basura porque es lo adecuado y genera consecuencias positivas tanto para mí como para la sociedad de la que formo parte, aunque no me penalicen ni me den ningún premio para hacerlo”), bajo el supuesto de que aquello que se hace por convicción difícilmente cambia, aunque cambien las condiciones externas (Tyler y Blader 2013). Los instrumentos de tipo *nudge*, por su parte, apoyados en los descubrimientos de la economía conductual y las neurociencias, son menos optimistas respecto de la racionalidad de los agentes, y apuntan a estructurar las alternativas de forma que “tomar la decisión” correcta sea lo menos costoso. La ausencia de saleros en las mesas o la donación de órganos como elección automática son ejemplos paradigmáticos de esta lógica<sup>4</sup>.

---

4 En el primer caso, ponerle sal a la comida no está prohibido ni supone pagar de más por la sal, pero sí implica el costo de pedir el salero en un lugar público, en el que la sanción social o las miradas de los comensales podrían representar un inhibidor. Asimismo, el que algunos países hayan definido que aquellas personas que al morir no quieran donar sus órganos tengan que hacer un trámite especial para explicitar su decisión, y que quien no lo haga demuestre su voluntad de donarlos de forma automática refleja muy bien esta lógica de

La evaluación de los instrumentos sustantivos supone analizar sus efectos. Estimar su idoneidad implica medir el grado de cambio generado por las acciones de política pública en los problemas que se quieren transformar. Sin embargo, la evaluación como instrumento orientado al aprendizaje debe ser transversal al proceso decisorio. En el siguiente apartado se hace hincapié y se desarrolla esta importante cuestión.

## La evaluación como una actividad transversal al proceso de políticas. Algunas claves de su uso y utilidad

Hasta aquí, el documento ha intentado analizar la importancia de la evaluación como (super)instrumento de política pública. Algo así como una meta-evaluación de la evaluación. Sin embargo, la evaluación tiene una importancia específica en cada una de las fases del ciclo de políticas (Banco Interamericano de Desarrollo [BID] 1997; Kellogg Foundation 1998; Lázaro 2015; OECD 2020). El objetivo, en este apartado, será reseñar los diversos criterios a partir de los cuales puede ser clasificada la actividad de evaluación y su utilidad para el análisis de políticas públicas a lo largo del *policy process*.

De acuerdo con el momento en que se realiza, la evaluación puede ser:

- *Ex ante*: se realiza antes de implementar el programa o política pública, y consiste en simular los efectos que causará en la población una determinada intervención estatal. A este tipo también se le conoce como de efectos anticipados. Su finalidad es decidir si debe o no implementarse la política pública (Cohen y Franco 2003).
- Intermedia: está orientada al análisis de la operación del programa durante su ejecución. Por lo general, adopta la forma de un control de gestión que permita hacer ajustes programáticos (en las metas, los recursos humanos o los procesos) en caso de ser necesarios.
- *Ex post*: se realiza una vez finalizada la intervención (Feinstein 2007) y permite emitir un juicio más general sobre sus resultados a partir de una reflexión comparativa entre lo que se planificó al inicio y los resultados e impactos reales que se obtuvieron (Ortegón, Pacheco y Prieto 2005).

Por su parte, según sus objetivos, las evaluaciones podrían dividirse en:

---

estructurar las decisiones para que sea más costoso decidir aquello que va en contra del bienestar social (Thaler y Sunstein, 2018).

- Diagnóstica: su finalidad es generar información sobre las características del problema público. Permite elaborar hipótesis sobre sus causas, diseñar alternativas de intervención e identificar a la población objetivo en cada caso;
- Formativa: su finalidad es detectar fallos o desvíos durante la implementación y corregir el rumbo cuando la situación así lo requiera;
- Sumativa: su finalidad es doble: a) conocer los resultados del programa y b) generar información para decidir qué hacer con él: mantenerlo, corregirlo o discontinuarlo.

Finalmente, de acuerdo con su diseño metodológico, las evaluaciones pueden tener las siguientes clasificaciones:

- Cuasi-experimentales, son métodos que asumen que las políticas públicas son tratamientos aplicados para el combate de un problema. En este tipo de diseños, los resultados de los programas se analizan a partir de la utilización de escenarios contrafactuales que permiten simular un experimento entre un grupo de tratamiento (el beneficiario del programa) y un grupo de control (igual en todo al de tratamiento excepto por el hecho de que no es beneficiario del programa). De esta forma, se puede establecer una relación causal entre los medios y las acciones implementadas por el programa y sus resultados; el grado de cambio logrado atribuible exclusivamente al tratamiento (Navarro 2005). Este tipo de evaluaciones utilizan métodos cuantitativos, específicamente la técnica de *matching*, en sus diversas versiones, es cada vez más utilizada para estimar el impacto neto de los programas sobre la población beneficiaria (King et al, 2007).<sup>5</sup>
- Observacionales; En estos casos, pueden utilizarse tanto métodos cuantitativos como cualitativos. En el primer caso, se utilizan bases de datos de N grande para estimar el resultados a partir de ciertas condiciones causales, pero la relación de causalidad entre las acciones del programa y los cambios experimentados por la población beneficiaria no pueden establecerse a partir de un escenario contrafactual; ya porque no se cuenta con la información adecuada para comparar a una comunidad que ha recibido apoyo del programa, frente a otra de similares características que no lo ha recibido (Navarro 2005), ya porque los objetivos de la evaluación son diferentes. Por su parte, las evaluaciones que adoptan una metodología observacional cualitativa, buscan -por lo general- asociar eventos con resultados mediante una lógica descriptiva que permite profundizar en el

---

5 Estas evaluaciones aíslan los efectos del programa del resto de los factores que pudieron haber influido en la situación de la población beneficiaria al momento de realizar la evaluación. Se recomienda revisar la página web de Gary King, para profundizar en el tema de los “matching methods” para el análisis de política pública. <https://gking.harvard.edu/category/research-interests/methods/causal-inference>

análisis de ciertas condiciones y también suelen ser muy útiles para dar voz a los actores involucrados, cuyos comportamientos adquieren sentido y brindan contexto a los resultados obtenidos mediante métodos cuantitativos (Pattyn, Molenveld y Befani, 2019; Ragin, 2007; Parsons 1995).

Tabla 1. El ciclo de políticas y las funciones de la evaluación

|                       | Fase del ciclo            | Actividad  | Pregunta que responde  | Función de la evaluación  |
|-----------------------|---------------------------|--|--|---|
| Evaluación ex ante    | Definición del problema   | Diagnóstico del problema público   | ¿Cuáles son las características del problema?  | Describir la magnitud del problema; identificar a la población afectada   |
|                       | Selección de alternativas | Ponderación y selección de alternativas de intervención a partir de criterios establecidos   | ¿Cuál es la mejor política dadas las restricciones existentes?   | Generar información para tomar la “mejor decisión”  |
| Evaluación intermedia | Diseño                    | Identificación de públicos objetivo, y diseño o selección de instrumentos de intervención  | ¿Qué acciones pudieran implementarse para prevenir/resolver el problema?   | Diseñar las reglas de operación del programa  |
|                       | Implementación            | Evaluación formativa o de “control de gestión”<br><br>(Procesos que permiten el seguimiento o monitoreo del funcionamiento del programa durante la ejecución)                | ¿El programa está operando de acuerdo a lo planeado? ¿Las acciones (metas) pensadas como medios para lograr los objetivos están siendo eficaces? ¿Cuál es el desempeño de los recursos humanos que operan el programa? | Permite decidir si mantener el programa tal como está operando o introducir ajustes para su mejora                                      |
| Evaluación post       | Terminación               | Evaluación sumativa. Se realiza al finalizar el programa y puede tener distintos objetivos (medir el impacto, conocer cómo operaron los procesos durante la ejecución, etc.) | ¿Cuáles fueron los resultados del programa? ¿Se alcanzaron las metas? ¿Se cumplieron los objetivos planteados? ¿A qué costo?   | Permite decidir si mantener o no el programa. Arroja información que puede servir de diagnóstico para una nueva etapa de implementación |

Elaborado por el autor.

Nota: En la literatura especializada es posible encontrar cierto desacuerdo respecto de los términos con que se clasifican a los tipos de evaluación en función de la fase del ciclo de políticas en que son introducidas. Mientras que para Cohen y Franco (2003, 110), la evaluación intermedia (o de control de gestión) es un tipo de evaluación ex post (porque se da posteriormente al inicio de la política) para Feinstein (2007, 20), la evaluación ex post es aquella que se da una vez pasada la fecha de finalización del programa.

Ubicar a los distintos tipos de evaluación con relación a la etapa del proceso de políticas en que es introducida puede resultar pertinente puesto que no solo se

identifica la dimensión temporal (el momento en que es ejecutada la evaluación), también se señalan los alcances y las metodologías específicas, utilizadas por cada evaluación. De hecho, en cada una de ellas los objetivos son diferentes y ayudan a responder distintos interrogantes.

Si analizamos las etapas con mayor detenimiento, podemos entender mejor la importancia de la evaluación –así como sus distintas versiones– a lo largo de todo el ciclo de una política pública.

### *La definición del problema*

En este punto, es importante tomar en cuenta que no todas las situaciones que afectan a una comunidad son consideradas un problema; es decir, no existen problemas por sí mismos, fuera del sujeto “problematizador”. Por tanto, la identificación y definición de una situación como problema ya implica cierta “evaluación”. La existencia de un problema remite a cierto juicio respecto de una situación que es anómala, un acontecimiento que no debería de ocurrir de acuerdo a determinados valores, saberes, sentidos comunes o perspectivas.

Ahora bien, ¿qué tan importante es el problema? ¿Es un problema público? ¿Cuál es la población afectada? ¿Debe el Estado, a través de sus poderes públicos, intervenir para su prevención y/o resolución? ¿Están dadas las condiciones para una intervención efectiva? ¿Se cuenta con los recursos –humanos, administrativos y financieros– para llevar adelante la intervención?

Las respuestas a estas preguntas hacen referencia a distintos modos de ponderación. Evaluar la “importancia del problema” y su nivel de “publicidad” remite a una discusión axiológica, en la que se contraponen modelos de sociedad, valores e ideologías. Lejos de la neutralidad valorativa, el proceso de definición del problema pone en contradicción enfoques parciales de una misma realidad.

Decidir si el problema es o no público, y si requiere o no de una acción colectiva para su solución, requiere de evidencia; para emitir un juicio se requiere de información. Por lo tanto, una decisión racional requiere de un diagnóstico que describa la magnitud del problema (la cantidad y tipo de población afectada) y analice la gravedad de sus consecuencias o sus implicaciones políticas).

## *La evaluación diagnóstica: descripción, estimación y selección de alternativas*

Una vez establecida la existencia de un problema público, conocidas su magnitud y sus características, así como sus posibles consecuencias, los funcionarios –con o sin el concurso de otros actores sociales– emprenden la tarea de diseñar alternativas de acción para resolver el problema en cuestión.

El primer paso en el diseño de una intervención pública es la definición de los objetivos que el proyecto se propone lograr. El segundo paso es determinar una serie de metas que deben alcanzarse como paso previo para el logro de los objetivos. Las metas (medios para el alcance de fines u objetivos), a su vez, deben observar una serie de criterios que guían la implementación; estos criterios están determinados en buena medida por las condiciones de factibilidad de la política. Finalmente, los criterios reflejan los valores que están detrás de las distintas alternativas de intervención sobre el problema definido como tal.

Revisemos un ejemplo. Un conjunto de actores sociales (activistas ambientales, el gobierno, organizaciones no gubernamentales, organismos internacionales) han identificado como problema público al cambio climático. A raíz de ello, la secretaria de medioambiente ha tomado la decisión de diseñar un programa público para atender este problema. El objetivo del programa es reducir en un 40 % de aquí a cinco años el número de emisiones de gases de efecto invernadero. Se solicita a dos grupos de trabajo en el interior de dicha secretaria presentar un curso de acción para el logro de dicho objetivo. Las metas seleccionadas por cada uno de los grupos son las siguientes: a) subsidiar el precio de las gasolinas sin plomo, b) ampliar el número de días por semana en que determinados vehículos no puedan circular.

**Pregunta:** ¿Es posible responder en el vacío cuál de los proyectos debería elegirse?

**Respuesta:** Para poder elegir uno de los dos proyectos deberían esgrimirse ciertos argumentos y justificar la pertinencia de la acción seleccionada. La decisión por tanto no puede tomarse sin evaluar cada uno de los dos cursos de acción de acuerdo con determinados criterios. ¿Qué política es más factible de ser implementada? ¿Qué política es más efectiva? ¿Cuál de las dos es menos costosa en términos financieros? Si el criterio de decisión es la eficiencia medida en términos económico-financieros, la política seleccionada debería ser la b), puesto que si bien existe un costo administrativo para el monitoreo del cumplimiento y la aplicación de las sanciones a los infractores, es de esperar que el monto de estas últimas sumado al beneficio social de la reducción de emisiones, excedan con creces al primero. Por el

contrario, si el criterio de decisión es la factibilidad política (es decir, la probabilidad de que la medida no despierte rechazo por parte de la sociedad en su conjunto), el curso de acción a elegir debería ser el a), ya que muy probablemente el proyecto b) generará un amplio rechazo entre los ciudadanos dada la arraigada “cultura de uso del automóvil” y la baja calidad del transporte público.

**Conclusión preliminar:** El ejemplo anterior –muy elemental– permite entender por qué la evaluación está presente desde un inicio en el ciclo de formulación de una política pública. Para tomar una decisión, no solo es necesario conocer a fondo el problema sobre el que se pretende intervenir, sino analizar y ponderar (a partir de una serie de criterios excluyentes entre sí) los costos y beneficios de cada una de las alternativas de acción disponibles.

Tabla 2. Ejemplos de preguntas diagnósticas

**Evaluación previa a la toma de decisiones o *ex ante***

¿Cuál es la magnitud del problema (qué tan grande o grave es)?

¿Cuáles son sus características?

¿Quiénes son los actores involucrados?

¿Cuáles son las causas del problema?

¿Cuáles son las alternativas de solución?

¿Qué instrumentos de política pública supone cada una de ellas?

¿Cuáles son los comportamientos que se esperan cambiar?

¿Qué bienes o servicios deben proveerse en cada caso?

¿Quiénes se benefician y quiénes se perjudican con cada una?

¿Cuáles y cuántos son los recursos necesarios para su puesta en marcha?

¿Cuál es la duración estimada de la intervención?

¿Cuál es la población objetivo del programa?

¿Cuál es la factibilidad (política, financiera, administrativa, etc.) de cada alternativa?

Fuente: Elaborado con base en Rossi, Lipsey y Freeman (2004).

En resumen, la evaluación *ex ante* (que sustenta la toma de decisiones) es un proceso sistemático de recogida y análisis de datos. Su objetivo es generar evidencia que permita comprender y describir los problemas a resolver. Constituye una hipótesis sobre las causas de un problema y, a partir de la adopción de determinados criterios, define cuál es la intervención más adecuada para su solución.

### *La evaluación formativa (durante la implementación)*

La implementación de un proyecto social implica un conjunto de operaciones limitadas en el tiempo, de las cuales resulta un producto final. La racionalidad de las políticas públicas indica que dichas intervenciones son actividades planificadas orientadas a la consecución de objetivos –previamente establecidos–. La ejecución de estos proyectos, su operación concreta a partir de una serie de restricciones (financieras, humanas, administrativas, sociales y organizacionales) es lo que se conoce como fase de implementación (BID 1997).

La evaluación intermedia o formativa es aquella que se realiza durante la operación del programa. Por lo general, este seguimiento o monitoreo suele ser muy útil puesto que ayuda a introducir ajustes “a tiempo”, cuando el rumbo del programa no es el originalmente planeado.

Paradójicamente, las evaluaciones formativas suelen ser las menos comunes. La vorágine de la operación, la poca disponibilidad de tiempo y recursos para “hacer un corte de caja” y seguir adelante, suelen atentar contra este tipo de evaluación. Por otro lado, para los funcionarios comprometidos con la implementación, suele ser muy difícil distanciarse razonablemente del proceso a fin de verificar que tan eficiente y efectivo está siendo este último. Los manuales de evaluación (BID 1997; Kellogg Foundation 1998), de hecho, sugieren que las evaluaciones de seguimiento o monitoreo sean realizadas por observadores externos, dada la necesaria distancia que debe haber entre el “sujeto evaluador” y el “objeto evaluado”, en este caso el programa.

Una dificultad adicional que deben enfrentar las evaluaciones de procesos es que constituyen muy buenas oportunidades para evaluar no solo la ejecución del programa, sino también el desempeño de quienes participan en la operación. Detrás de la evaluación centrada en la gente, está la creencia de que se necesita más control sobre los burócratas para garantizar que los programas diseñados por los hacedores sean implementados de manera eficiente y efectiva (Parsons 1995, 561).

Fue así como, pese a las resistencias, la aplicación de múltiples indicadores de desempeño en el interior de la administración pública (calidad total, orientación al



cliente, presupuesto por resultados, incentivos por desempeño) se ha generalizado. De la mano de los enfoques gerenciales propios de la nueva gestión pública, se han plasmado distintos programas e incentivos para que quienes se dedican a la administración pública sean más competentes, tengan un mayor compromiso con su trabajo y desarrollen las habilidades necesarias para lograr que las políticas obtengan mejores resultados.<sup>6</sup>

Tabla 3. Ejemplos de preguntas sobre la operación del programa

### Evaluación del proceso del programa

- ¿Se están alcanzando las metas del programa?
- ¿Las metas del programa están contribuyendo a alcanzar los objetivos?
- ¿La población beneficiaria es un subconjunto de la población objetivo?
- ¿Cuál es el porcentaje de la población objetivo que se está atendiendo?
- ¿Quiénes reciben los servicios del programa lo hacen de forma completa?
- ¿Hay diferencias entre la estimación de los recursos a invertir y la inversión real?
- ¿Los supuestos originales se están verificando?
- ¿Se está cumpliendo con las cláusulas del contrato, convenio o reglas de operación?
- ¿Los recursos financieros se han utilizado de forma eficiente?
- ¿Los recursos humanos del programa cumplen con sus responsabilidades?
- ¿Los recursos humanos son suficientes?
- ¿Es necesario introducir ajustes en el rumbo del programa? ¿Es posible?

Fuente: Elaborado con base en Rossi, Lipsey y Freeman (2004).

Resumiendo, las evaluaciones formativas o intermedias forman parte de lo que se conoce como “control de gestión”. Los objetivos del control son asegurar la capacidad de ejecución y garantizar la conformidad de esta con el rumbo marcado en la planeación. Ello exige verificar sistemáticamente lo realizado respecto de lo planeado, corregir oportunamente los desvíos y reprogramar o mejorar las metas establecidas originalmente (BID 1997). Así,

6 Salvo en el caso de la evaluación de procesos por medio de la cual pueden detectarse fallos de implementación (Weiss 1998), la mayor parte de la literatura dedicada a evaluar a los recursos humanos forma parte de lo que se conoce como nueva gestión pública (Parsons 1995).

pese a no estar muy arraigadas ni ser muy populares entre los administradores públicos, este tipo de evaluaciones constituyen un elemento clave para el éxito del proyecto.

### *La evaluación “sumativa” o ex post: un análisis de los resultados del programa*

Tradicionalmente, la evaluación como actividad está asociada a aquellos juicios realizados al final de determinado proceso. La evaluación tiene lugar, desde esta perspectiva, cuando necesitamos aprender de la experiencia pasada, conocer qué tanto se logró de lo que nos propusimos, a qué costo y cuáles fueron los elementos que no funcionaron de acuerdo con lo esperado (errores de cálculo o de procedimiento).

La evaluación “sumativa” o *ex post* es aquella que se realiza con posterioridad a la fecha de terminación de un proyecto (Feinstein 2007). Este tipo de evaluación suele tener dos objetivos principales, uno retrospectivo y otro prospectivo. En términos retrospectivos, la evaluación “sumativa” analiza, a grandes rasgos, cuáles fueron los resultados de la intervención pública sobre el problema en cuestión. En términos prospectivos, permite aprender de la experiencia extrayendo lecciones a futuro sobre lo que ha funcionado bien y lo que ha fallado, pero fundamentalmente ofrece un nuevo diagnóstico del problema abordado por la política pública (Tabla 4).

La evaluación de impacto analiza cuál ha sido la magnitud del cambio en el problema que es atribuible exclusivamente a la intervención. En tal sentido, lo que se evalúa aquí es el grado de alcance de los objetivos planteados por el proyecto o programa, es decir, su efectividad (Cohen y Franco 2003).

Con la evaluación o el análisis del costo-beneficio se intenta conocer cuál fue el beneficio del proyecto respecto de los costos, en términos monetarios. Esta metodología se interesa por cuantificar las ganancias netas obtenidas mediante el proyecto. Para el caso de los proyectos sociales, el análisis costo-beneficio debe incorporar los costos sociales de la intervención, también conocidos en la literatura especializada como “precios sombra”. El “precio sombra” no implica una carga financiera para el proyecto, pero sí representan un costo para la sociedad, motivo por el cual deben de cuantificarse —a partir de simulaciones— y añadirse al cálculo final del beneficio neto del proyecto (Boardman et al. [1996] 2001). En este sentido, los impactos ambientales de los proyectos de infraestructura son un excelente ejemplo. En los casos de construcciones de carreteras, el análisis costo-beneficio debe contemplar no solo el costo financiero de su construcción; desde un enfoque de política

pública debe de poder cuantificarse cuánto “pagará” la sociedad a raíz del impacto ambiental de la tala de bosques, la pérdida de ecosistemas, etc. El criterio que está detrás de este análisis suele ser el de eficiencia o eficiencia social.

Tabla 4. Ejemplos de preguntas sobre los resultados del programa

**Evaluación *ex post***

- ¿Se han cumplido las metas del programa? (Eficacia)
- ¿Las metas del programa han contribuido a alcanzar los objetivos?
- ¿En qué medida se alcanzaron los objetivos del programa? (Efectividad)
- ¿Cuáles fueron los efectos de la política sobre la población beneficiaria del programa?
- ¿Hubo efectos no deseados como consecuencia de la implementación del programa?
- ¿Cuál es el nivel de satisfacción de la población objetivo con el programa?
- ¿Dados los recursos disponibles, se ha cumplido con las metas establecidas al mínimo costo? (Eficiencia)
- ¿Cuál ha sido el beneficio neto del programa en términos monetarios? (Relación costo-beneficio)
- ¿Cuál ha sido la magnitud del cambio en el problema público que puede atribuirse exclusivamente a la intervención? (Impacto)
- ¿Dados los recursos invertidos, se ha maximizado el impacto del programa? (Costo-efectividad)
- ¿Qué hubiera pasado si el programa no se hubiera implementado?

*Fuente:* Elaborado con base en Cohen y Franco (2003).

***Tipos de evaluación sumativa o ex post***

La evaluación de procesos focaliza su atención en los mecanismos e instrumentos a través de los cuales se ha intentado alcanzar los objetivos del programa o proyecto. Una vez realizada la evaluación de impacto o el análisis del costo-beneficio, se dispone de una cantidad de información que permite hacer inferencias sobre aquellas cuestiones del programa que no funcionaron según lo esperado; cuando se descarta la posibilidad de errores en el diseño, dichas inferencias centran su mirada en la fase de implementación, y es allí –en el proceso de operación del programa– donde se posa la mirada de los evaluadores. El criterio determinante de este tipo de evaluación es la eficacia;

es decir, el haber cumplido con las metas establecidas en el diseño como medio para el logro de los objetivos (Feinstein 2007)<sup>7</sup>.

## Los criterios de la evaluación. De las ideas a la evidencia

Los criterios de una evaluación son aquellos valores que sustentan las acciones de las políticas; son de alguna forma los principios que justifican la intervención y sus particularidades. Los criterios de evaluación constituyen, por tanto, la base de la toma de decisiones. Están enfocados en fines y, por eso, cada uno de ellos aborda diferentes aspectos de un problema público (equidad, sustentabilidad, respeto por la diversidad) o de la política en cuestión (eficiencia, efectividad, transparencia, sostenibilidad financiera). De ahí que una política puede ser evaluada a partir de sus resultados (cuáles, y de qué magnitud han sido sus efectos sobre el problema), sus operaciones y procedimientos (de qué modo ha logrado tales resultados).

Existen muchos criterios posibles de evaluación de una política pública: tantos como los fines que se proponga alcanzar. En sentido estricto, los criterios son mutuamente excluyentes. Sin embargo, una política puede estar orientada a cumplir con más de un criterio y estos últimos pueden ser localizados a lo largo de todo el *policy process*.

### a) Criterios que guían la selección de alternativas

- Factibilidad jurídica: ¿el proyecto, tal y como fue diseñado, viola alguna normatividad?
- Sostenibilidad financiera: ¿los recursos para llevar adelante el proyecto están disponibles o se pueden conseguir?
- Factibilidad política: ¿hay actores sociales o políticos con capacidad como para impedir la puesta en marcha del proyecto?
- Sustentabilidad de los resultados: una vez finalizado el proyecto, ¿los resultados podrán mantenerse en el tiempo sin que sea necesaria una nueva intervención?
- Capacidad administrativa: ¿están disponibles los recursos humanos y organizativos para que el proyecto pueda ser operado siguiendo la planeación?

---

7 Si la evaluación de impacto demostró que no se cumplió con los objetivos tal y como se había calculado, y la evaluación de procesos posterior mostró que el programa había sido eficaz en el cumplimiento de las metas y los procesos establecidos, es posible inferir que el error fue de diseño; es decir, de estimación o planeación. Muy probablemente, las metas establecidas no eran los mejores medios para lograr los objetivos, o se calculó mal el tiempo que debía durar el proyecto, o peor aún, la selección de la población objetivo (sea la identificación o la selección de beneficiarios propiamente dicha) fue equivocada.

- Costo-beneficio/eficiencia: ¿manteniendo constantes los costos, el proyecto maximiza el bienestar, medido en términos monetarios? ¿Dado el beneficio que se pretende obtener, el proyecto minimiza los costos? Este análisis es agregado y considera el beneficio sobre la sociedad en su conjunto (sin importar como se distribuyen los costos) en cuanto base para la decisión.
- Costo-impacto/efectividad: ¿Manteniendo constantes los costos, el proyecto maximiza los efectos positivos sobre el problema a tratar? ¿Dado el impacto que se busca alcanzar, el proyecto minimiza los costos? Se considera el impacto sobre la población objetivo, fijada según los objetivos del proyecto.

b) Criterios que guían la operación del programa

- Eficacia: es el grado en que se realizan las acciones y se alcanzan las metas establecidas en la planeación.
- Eficiencia: es la medida en que se realizan las acciones y se alcanzan las metas al menor costo posible.
- Efectividad: es el grado en que el proyecto alcanza sus objetivos.
- Orientación al cliente/ciudadano: ¿este proyecto incorpora las expectativas o preferencias de la población objetivo?; este criterio es más cualitativo y si bien no prescinde de las consideraciones relativas a los resultados del programa, utiliza indicadores que miden el grado de satisfacción de los beneficiarios con la intervención en su conjunto.

c) Criterios que guían el análisis de los resultados

- En las evaluaciones sumativas, realizadas una vez finalizado el programa o proyecto, se analizan los resultados del proyecto en función de lo programado originalmente. Estos parámetros pueden ser metas (eficacia), beneficios netos (eficiencia), objetivos (efectividad), satisfacción del ciudadano, etc. En estricto sentido, muchos de los criterios ya analizados pueden utilizarse para evaluar la política implementada. Lo importante es que, para cada pregunta, existe un tipo de evaluación diferente. Para evaluar el alcance de las metas, se realizan análisis de costo-eficacia; para conocer los beneficios netos de un programa, se analiza la relación costo-beneficio; para obtener una medida del cambio logrado por la intervención en la situación definida como problema, se realizan análisis de costo-efectividad o costo-impacto (ACI) (Cohen y Franco 2003). En todos estos casos, los diseños son cuasi-experimentales y debe contarse con datos de la situación del problema antes y después de la implementación del programa.

## ¿Evidence-based policies vs politics-based policies?: La evidencia y la política como fuentes de decisión

En un enfoque “tecnocrático” de *policy analysis*, se considera que las políticas públicas son formuladas por planificadores racionales y benevolentes (Blankart y Koester 2006). Todo aquel interesado en alcanzar mejores resultados sociales tendría simplemente que identificar situaciones problemáticas, examinar las alternativas de acción, evaluar cuáles son las que maximizaran los efectos positivos (impacto) y seleccionar aquella más factible dadas las restricciones jurídicas, financieras y administrativas.

Esta perspectiva parte de dos supuestos, uno correcto y otro erróneo. El correcto: considerar a la construcción de la evidencia la base de la toma de decisiones racional. El equivocado: considerar a la política (los intereses en disputa, los procesos de negociación, el liderazgo político, la deliberación pública y el marco institucional con moldea las interacciones entre estos) un factor exógeno, externo al proceso formulación de políticas públicas.

Una política basada en la evidencia utiliza consciente, explícita y juiciosa la “mejor” evidencia científica disponible para la toma de decisiones (Pinilla Pallejá 2006). La evidencia que soporta el diseño de políticas públicas procede de la investigación básica o aplicada, de acuerdo con el ámbito de intervención. Este enfoque emerge en contraste con las “políticas basadas en la opinión”, fuertemente sustentadas en la utilización “selectiva” (por no decir “sesgada”) de testimonios de grupos o individuos sustentados a su vez en puntos de vista ideológicos, prejuicios o conjeturas especulativas (Davies 2004).

Sin embargo, si la evidencia científica es imprescindible para la toma de decisiones, no lo es menos el hecho de que los factores políticos constituyen una parte importante de esa información. Por ello, se parte aquí de la premisa según la cual los procesos de debate, negociación, aprobación y ejecución de políticas pueden ser al menos tan importantes como los contenidos específicos de las mismas (Del Tronco 2018).

Con respecto a la evaluación, solo unos cuantos autores próximos a enfoques cualitativos hubieran suscrito, hasta hace poco tiempo, que dicha actividad constituye un proceso eminentemente político (MacDonald 1976). Desde hace unos años, y especialmente en América Latina, este enfoque ha conseguido un creciente número de seguidores en el seno del *policy analysis*.

Algunos de los argumentos esgrimidos a favor de esta postura se exponen a continuación.

Las evaluaciones tienen como objeto de análisis programas y proyectos que son propuestos, formulados, debatidos, financiados e implementados en el seno de un proceso político (Palumbo 1987). Los hacedores y especialmente los funcionarios

implicados en la política pública evaluada inciden notablemente en su realización (García Sánchez 2009, 129).

El impulso de las actividades de evaluación procede, fundamentalmente, de los poderes públicos, en forma de institucionalización o de financiación. Los gobiernos ven en la evaluación un instrumento para legitimar la orientación de sus políticas tanto como una oportunidad para el control de la administración. Como afirma House (1980), el gobierno legitima la evaluación y la evaluación legitima la actuación gubernamental (García Sánchez 2009, 129).

Por otro lado, la evaluación pone en discusión la pertinencia de determinados criterios a partir de los cuales dicha política es ponderada y enjuiciada. El conjunto de criterios utilizados no siempre es exhaustivo y, aunque lo fuera, la jerarquización que se hace de cada uno de ellos responde, por lo general, a cuestiones políticas cuando no partidistas. Así, en el centro de esta discusión las consideraciones “políticas” (*politics*) pueden llegar a sobrepasar a las consideraciones “de política” (*policy*).

Desde este abordaje “político” de las políticas públicas, es posible explicar por qué determinadas reformas son exitosas en países de una misma región, pero fallan en otros, o en países que comparten una cultura común, pero cuyos contextos político institucionales y arreglos de economía política son diferentes. De tal manera, los resultados de las políticas pueden explicarse también por la naturaleza del sistema de partidos, la estructura y funcionamiento de la legislatura, los incentivos institucionales del presidente o la normatividad que rige a la administración pública, y por la validez de la teoría de cambio que estructura y sostiene el diseño metodológico de la evaluación.<sup>8</sup>

### *La deliberación pública. ¿Una estrategia de reconciliación de ambos enfoques para la evaluación de políticas?*

El debate tradicional entre tecnócratas y demócratas gira en torno a si las políticas públicas deber ser el resultado de procesos de negociación política, en los que se imponga la postura de los actores mayoritarios o más poderosos, o deben ser, más bien, la consecuencia de un análisis razonado de la evidencia científica disponible. Más arriba, se presentaron argumentos para justificar la incorporación de ambos criterios, la legitimidad democrática y la racionalidad técnica, en el proceso decisorio (Del Tronco 2020, 47). La pregunta clave es cómo garantizarlo.

---

8 La institucionalización del sistema de partidos (es decir, la capacidad que tienen los partidos de mantenerse en el tiempo como alternativas de elección popular) influye en la definición de una agenda de gobierno, en las posibilidades de coordinación Ejecutivo-Legislativo y en los incentivos de los funcionarios electos para atender conjuntos más estrechos o más amplios de intereses en la sociedad. Finalmente, todo ello incide en el grado de satisfacción de la ciudadanía con el desempeño gubernamental (Stein y Tommassi 2006, 29-31).

Siguiendo a Majone (1997), uno de los aportes más útiles que pueden hacer los defensores de la deliberación como mecanismo de toma de decisiones es ayudar a los gobernantes y a la opinión pública a evitar que se exageren los límites de lo posible en la política. En un análisis deliberativo (tecnopolítico) de las políticas se trata de identificar todas las restricciones reales o potenciales enfrentadas por el gobierno, separarlas de los obstáculos ficticios, evaluar su importancia para diferentes estrategias de ejecución y estimar los costos y beneficios que implica relajar aquellas restricciones que no sean necesariamente fijas (Majone 1997, 109).

Desafortunadamente, dice Majone, el análisis de factibilidad no siempre se incorpora a los procesos de evaluación *ex ante*. Los analistas de política se ocupan de manera explícita de las restricciones técnicas o presupuestarias, pero tienden a subestimar las limitantes políticas e institucionales, cuando no hay ninguna razón para ello: todas restringen la libertad de acción del decisor, y su violación lleva consigo siempre una sanción (Majone 1997, 110). Tales restricciones explican por qué es tan improbable –en política pública– implementar una decisión óptima, y hay que contentarse con políticas de *second best* (segunda mejor opción) u “óptimo condicionado”. De acuerdo con Majone, si estas políticas son las únicas viables, es la factibilidad y no lo óptimo el criterio principal que orienta la decisión.

Es cierto que hay ámbitos de política pública que por tradición (o necesidad) han estado más conectados con (o colonizados por) el mundo académico o tecnocrático (Löfgren y Bickerton 2013). La política económica, la política energética, la política de empleo y, en algunos países, también las políticas ambientales o de salud han estado más expuestas al estado del conocimiento y en particular a las corrientes dominantes dentro de las ciencias económicas, cuyos paradigmas suelen gobernar los procesos de formulación, diseño y evaluación de las políticas públicas.

Asimismo, los analistas han advertido que la prestación efectiva de los servicios públicos requiere de algo más que el diseño de un óptimo, desde el punto de vista teórico. Tan importante como esto último resultan la forma en que se aplica el programa y la detección de quiénes se benefician y quiénes se perjudican con la intervención, si el programa está logrando lo que se quería y, en caso contrario, cómo podrá mejorarse o abandonarse.

En tal sentido, es importante subrayar que cuando hablamos de conocimiento no siempre reparamos en que dicho término tiene diversas acepciones o dimensiones. El conocimiento epistémico es el que hace referencia a la evidencia científica producida a través del método hipotético-deductivo. El conocimiento técnico, por su parte, es aquel derivado de la aplicación práctica (técnica) de los hallazgos alcanzados por la ciencia. En tanto, el conocimiento normativo o valorativo es aquel que se deriva de las ideas de lo adeudado compartidas por una comunidad. De acuerdo



con Majone (1997), las políticas más efectivas son aquellas que incorporan las tres formas de conocimiento, pero suele ser muy poco usual. Para alcanzarlo de forma habitual, sería necesaria la interacción de actores con capacidad de producir y aplicar las diversas formas de conocimiento, y ello requeriría de procesos de formulación y diseño participativos, en los que las ideas, la práctica y la evidencia fueran consideradas (Löfgren y Bickerton 2013).

Para todo ello, los espacios deliberativos, contruidos a partir de procesos de diálogo incluyentes compuestos por una muestra representativa de los actores involucrados suele ser una buena –si no la más adecuada– estrategia de intervención (Fishkin 2003; Christiano 2012). Hablar de racionalidad en las políticas debería suponer, entonces, la combinación de conocimientos técnico-científicos y consensos normativos. Para decirlo en términos “weberianos”, la racionalidad debe ser tanto instrumental como valorativa.

El desafío resultante para la evaluación de políticas públicas, sin embargo, no es menor. ¿Cómo incorporar en los modelos la ponderación del impacto técnico como la posibilidad de medir los niveles de satisfacción de los *stakeholders*?

## Conclusiones.

### Mitos y realidades de la evaluación de políticas

Pese al velo de rigurosidad científica que viste a la evaluación de políticas públicas, gran parte del funcionariado la considera una actividad poco útil y en ciertos casos peligrosa. En una de las mayores críticas se señala que los métodos científicos de evaluación generan una enorme cantidad de datos, pero muy pocas novedades a la hora de las conclusiones (McNamara 2002). Pese a su sofisticación, los reportes de evaluación tienden a reiterar lo obvio, hecho que –cuando no cuestiona la actividad de los tomadores de decisión– refuerza el escepticismo de estos últimos sobre la utilidad de los primeros.

Otro de los mitos acerca de la evaluación sostiene que evaluar implica determinar el éxito o el fracaso de un programa público. Con esta postura se asume que un programa exitoso es el que resuelve el problema y satisface plenamente los intereses de todas las personas involucradas. Su uso común entre los evaluadores expertos explica por qué las actividades de evaluación son percibidas como una amenaza y los evaluadores como inquisidores dispuestos a quemar en la hoguera a quienes no formulen, diseñen y operen los programas de acuerdo con determinados criterios.

El éxito en todo caso significa que el programa permanezca abierto a una actividad de crítica, al ajuste permanente, de forma que pueda adaptarse a los cambios

del entorno y de los actores participantes. Corregir aquello que no funcionó a partir del aprendizaje que supone la evaluación es lo que termina de dar sentido a este instrumento y, por tanto, no puede haber política exitosa sin evaluación, pero el éxito de esta última depende también de una sofisticación metodológica que incorpore en la modelación los criterios tanto técnicos como políticos.

## Referencias

- BID (Banco Interamericano de Desarrollo). 1997. *La evaluación: una herramienta de gestión para mejorar el desempeño de los proyectos*. Washington D.C.: BID.
- Bemelmans-Videc, Marie Louise. 1998. "Introduction: Policy Instrument Choice and Evaluation". En *Carrots, Sticks and Sermons. Policy instruments and their evaluation*, editado por Marie Louise Bemelmans-Videc, Ray C. Rist y Evert Vedung, 1-18. Nueva Jersey: Transaction Publishers.
- Blankart, Charles, y Gerrit Koester. 2006. "Political Economics versus Public Choice". *KYKLOS International Review for Social Sciences* 59 (2): 171-200. <https://bit.ly/3fDRyDM>
- Boardman, Anthony E., David H. Greenberg, Aidan R. Vining y David L. Weimer. (1996) 2001. *Cost-Benefit Analysis: Concepts and Practice*, 2.<sup>a</sup> ed. Nueva Jersey: Prentice Hall.
- Christiano, Thomas. 2012. "Rational Deliberation Among Experts and Citizens". En *Deliberative Systems: Deliberative Democracy at the Large Scale*, editado por John Parkinson y Jane Mansbridge, 27-51. Cambridge: Cambridge University Press.
- Cohen, Ernesto, y Rolando Franco. 2003. *Evaluación social de proyectos*. Ciudad de México: Siglo XXI.
- Davies, Philip. 2004. "Is Evidence-Based government possible?". Jerry Lee Lecture presentada en el Annual Campbell Collaboration Colloquium, Washington D.C., 19 de febrero.
- Del Tronco, José. 2018. *La negociación efectiva y el proceso representativo. Apuntes para mejorar la formulación de políticas en democracia*. Ciudad de México: Instituto Electoral de la Ciudad de México.
- 2020. "Las ideas y las políticas públicas". En *Enfoques teóricos de políticas públicas: desarrollos contemporáneos para América Latina* en del Castillo, coordinado por Gloria Del Castillo Alemán y Mauricio I. Dussauge Laguna, 25-54. Ciudad de México: FLACSO México.

- Feinstein, Osvaldo. 2007. "Evaluación pragmática de políticas públicas". *ICE: Revista de Economía* 837: 19-32. <https://bit.ly/45tt09s>
- Fishkin, James S. 2003. "The Voice of the People". En *Democracy Sourcebook*, editado por Robert A. Dahl, Ian Shapiro y José Aantonio Cheibub, 25-28. Londres: Massachusetts Institute of Technology Press.
- Fontaine, Guillaume. 2015. *Análisis de políticas públicas. Conceptos, teorías y métodos*. Anthropos / Quito: FLACSO Ecuador.
- García Sánchez, Ester. 2009. "Metaevaluaciones de políticas públicas: una visión desde la ciencia política". *Revista del CLAD Reforma y Democracia* 43: 127-154. <https://bit.ly/41RiYf9>
- Hood, Christopher. 1986. "Exploring Government's Toolshed". En *The Tools of Government*, 1-15. Londres: Chatham House Publishers.
- House, Ernest. R. 1980. *Evaluating with validity*. Beverly Hills: SAGE.
- Howlett, Michael P. 2011. *Designing Public Policies. Principles and Instruments*. Londres: Taylor & Francis Group.
- Jerit, Jennifer. 2008. "Issue Framing and Engagement. Rhetorical Strategy in Public Policy Debates". *Political Behavior* 30: 1-24. <https://doi.org/10.1007/s11109-007-9041-x>
- Kellogg Foundation. 1998. "Manual de evaluación". Acceso el 5 de septiembre de 2014. <https://bit.ly/3MjEXWw>
- King, Gary et al. 2007. "A Politically Robust Experimental Design for Public Policy Evaluation, with Application to the Mexican Universal Health Insurance Program", en *Journal of Policy Analysis and Management*, Vol. 26, No. 3, 479-506. Published online in Wiley InterScience ([www.interscience.wiley.com](http://www.interscience.wiley.com)) DOI: 10.1002/pam.20279
- Knoepfel, Peter, Corinne Larrue, Frédéric Varone y Miriam Hinojosa Dieck. 2007. "Hacia un modelo de políticas públicas operativo". *Ciencia Política* 3: 6-29. <https://bit.ly/41N0nAU>
- Lascoumes, Pierre, y Legalés, Patrick. 2019. "Comprensión de las políticas públicas a través de instrumentos. De la naturaleza de los instrumentos a la sociología de la instrumentación en políticas públicas". En *Los instrumentos de política pública*, compilado por Eduardo Villarreal, 211-242. Ciudad de México: Siglo XXI.
- Lázaro, Blanca. 2015. "Comparative study on the institutionalisation of evaluation in Europe and Latin America". Study 15, Eurosocial Programme. <https://bit.ly/43h5YRF>
- Löfgren, Karl, y Sara Hendrica Bickerton. 2013. "How do policy professionals in New Zealand use academic research in their work?". En *Learning Policy, Doing Policy, Interactions Between Public Policy Theory, Practice and Teaching*, editado

- por Trish Mercer, Russell Ayres, Brian Head y Jonh Wanna, 107-128. Cambera: Australian National University Press.
- MacDonald, Barry. 1976. Evaluation and the control of education. En D. Tawney (Ed.), *Curriculum evaluation today: trends and implications*, 125-136. London: McMillan.
- Majone, Giandomenico. 1997. *Evidencia, argumentación y persuasión en la formulación de políticas*. Ciudad de México: Fondo de Cultura Económica.
- March, James, y Johan Olsen. 2006. "The Logic of Appropriateness". En *Oxford Handbook of Public Policy*, editado por Robert Goodin, Michael Moran y Martin Rein, 689-708. Londres: Oxford University Press.
- McNamara, Carter. 2002. "Basic Guide to Program Evaluation". <https://bit.ly/3WhYD1r>
- Morales, Sabina, Marcela Meneses y Silvia Alonso. 2006. Índice de transparencia para programas sociales. Ciudad de México: Alianza Cívica.
- Navarro, Hugo. 2005. *Manual para la evaluación de impacto de proyectos y programas de lucha contra la pobreza*. Santiago de Chile: CEPAL.
- OECD (Organización para la Cooperación y el Desarrollo Económicos). 2020. "How Can Governments Leverage Policy Evaluation to Improve Evidence Informed Policy Making". <https://bit.ly/3MJV5Sg>
- Ortegón, Edgar, Juan Francisco Pacheco y Adriana Prieto. 2005. *Metodología del marco lógico para la planeación, seguimiento y evaluación de proyectos sociales*. Santiago de Chile: CEPAL.
- Page, Edward C. 2006. "The origins of policy". En *Oxford Handbook of Public Policy*, editado por Robert Goodin, Michael Moran y Martin Rein, 200-227. Londres: Oxford University Press.
- Palumbo, Dennis J. 1987. *The politics of program evaluation*. Beverly Hills: SAGE.
- Parsons, Wayne. 1995. *Public Policy. An introduction to the theory and practice of policy analysis*. Northampton: Edward Elgar.
- Patton, Michael Quinn. 2002. *Qualitative Research & Evaluation Methods*. 3.<sup>a</sup> ed. Londres / Nueva Delhi: SAGE.
- Pattyn, Valérie, Astrid Molenveld and Barbara Befani. 2019. "Qualitative Comparative Analysis as an Evaluation Tool: Lessons From an Application in Development Cooperation". *American Journal of Evaluation* 2019, Vol. 40(1) 55-74 DOI:10.1177/1098214017710502
- Pinilla Pallejá, Rafael. 2006. "Agencia de evaluación: innovación social basada en evidencia". Documento de trabajo 91, Fundación Alternativas. <https://bit.ly/3pRw6Uh>

- Pressman, Jeffrey y Aaron Wildawsky. 1998. *La implementación*. Ciudad de México: Fondo de Cultura Económica.
- Ragin, Charles. 2007. *La construcción de la investigación social. Introducción a sus métodos y diversidad*. Bogotá: Universidad de los Andes
- Rossi, Peter Henry, Mark Lipsey y Howard E. Freeman. 2004. *Evaluation. A systematic Approach*. Londres / Nueva Delhi: SAGE. <https://bit.ly/3BH-g9mf>
- Schedler, Andreas. 2008. “¿Qué es la rendición de cuentas?”. *Cuadernos de Transparencia 3*: 1-50. <https://bit.ly/3MImnIU>
- Stein, Ernesto, y Mariano Tommassi. 2006. *La política de las políticas públicas*. Ciudad de México: Editorial Planeta.
- Subirats, Joan. 2015. “Todo se mueve. Acción colectiva, acción conectiva. Movimientos, partidos e instituciones”. *Revista Española de Sociología 24*: 123-131. <https://bit.ly/3MGfdDZ>
- Subirats, Joan, y Ricard Gomà. 1989. “Evaluación de políticas públicas”. En *Análisis de políticas públicas y eficacia de la administración*, editado por Joan Subirats. Madrid: Instituto Nacional de Administración Pública.
- Sunstein, Cass, y Richard Thaler. 2018. *Un pequeño empujón (Nudge). El impulso que necesitas para tomar mejores decisiones sobre salud, dinero y felicidad*. Madrid: Taurus.
- Tyler, Tom R., y Steven L. Blader. 2013. *Cooperation in groups. Procedural Justice, Social Identity and Behavioral Engagement*. Filadelfia: Tylor and Francis
- Weiss, Carol H. 1998. *Evaluation*. Nueva Jersey: Prentice Hall.