

EVALUACIÓN DE POLÍTICAS PÚBLICAS. TEORÍA Y TÉCNICAS

JESÚS CARRILLO RODRÍGUEZ



FACULTAD DE DERECHO

EVALUACIÓN DE POLÍTICAS PÚBLICAS. TEORÍA Y TÉCNICAS

Jesús Carrillo Rodríguez



Carrillo Rodríguez, Jesús

Evaluación de políticas públicas : teoría y técnicas / Jesús Carrillo Rodríguez.— Primera edición.— Mexico : Universidad Anáhuac México. Facultad de Derecho, 2024.

1 recurso en línea (188 páginas).

Bibliografía al final de los capítulos

ISBN 978-607-8566-78-5

1. Política pública -- Evaluación. 2. Agencias administrativas -- Evaluación. 3. Administración pública.

L.C.

Dewey

JF1525.P6

320.6

C37

C37

2024

2024

Diseño de portada: VLA. Laboratorio Visual

Diseño de interiores: Nuria Saburit Solbes

Primera edición, 2024

ISBN: 978-607-8566-78-5

DOI: https://doi.org/10.36105/publicaciones_derecho.2024.01

La presente edición digital de la obra

Evaluación de políticas públicas. Teoría y técnicas

le pertenece al editor mediante licencia exclusiva. El editor autoriza el acceso a la totalidad de la obra para su consulta, reproducción, almacenamiento digital en cualquier dispositivo e impresión para su uso personal y privado y sin fines de lucro. Ninguna parte de la presente obra podrá ser alterada o modificada ni formar parte de nuevas obras, compilaciones o colecciones. Queda prohibida su difusión y comunicación pública en plataforma digital alguna distinta a la cual se encuentra almacenada, sin permiso previo del editor.

Este libro fue arbitrado y dictaminado positivamente por dos dictaminadores externos bajo el sistema doble ciego. Ha sido valorado y liberado para su publicación por el Comité Editorial de la Universidad Anáhuac México.

Derechos reservados:

© 2024, Investigaciones y Estudios Superiores SC

Universidad Anáhuac México

Av. Universidad Anáhuac 46, Col. Lomas Anáhuac

Huixquilucan, Estado de México, C.P. 52786

Miembro de la Cámara Nacional de la Industria Editorial Mexicana.

Registro núm. 3407

DIRECTORIO

DR. CIPRIANO SÁNCHEZ GARCÍA, L.C.
Rector

DRA. LORENA ROSALBA MARTÍNEZ VERDUZCO
Vicerrectora Académica

DR. JOSE RODRIGO POZÓN LÓPEZ
Vicerrector Académico

DR. JOSÉ HONORIO CÁRDENAS VIDAURRI
Director de Investigación

MTRA. ADRIANA SÁNCHEZ ESCALANTE
Editora de Publicaciones

Contenido

Presentación	12
CAPÍTULO 1. Conceptos y antecedentes	15
Introducción	15
1.1 Conceptos básicos	15
1.2 Orígenes de la evaluación	16
1.3 La evaluación y sus dominios	18
1.3.1 Ciencias administrativas	18
1.3.2 Desarrollo económico	20
1.3.3 Educación	22
1.3.4 Investigación	23
1.3.5 Políticas públicas	24
1.4 Síntesis capitular	27
Referencias	27
CAPÍTULO 2. Ámbitos, modalidades y propósitos	29
Introducción	29
2.1 Los escenarios de la evaluación	29
2.1.1 Condiciones favorables y aprendizaje	31
2.1.2 Las organizaciones en torno a la evaluación	31
2.2 Modos de evaluación	34
2.2.1 Según la finalidad	34
2.2.2 Según el momento	35
2.2.3 De acuerdo con sus enfoques	36
2.2.4 De acuerdo con la intensidad y profundidad	38
2.2.5 Según los actores	40
2.2.6 Según la adscripción del evaluador	41
2.3 Espíritu de la evaluación	42
2.4 Síntesis capitular	44
Referencias	44
CAPÍTULO 3. Proceso de evaluación	46
Introducción	46
3.1 Procedimiento general de la evaluación	46
3.2 El procedimiento de la evaluación	47

3.3	Reflexión adicional acerca de las evidencias en la evaluación	49
3.4	La estrategia en el diseño	51
3.4.1	Diseño estratégico de una evaluación	51
3.4.2	Tipos de diseños de evaluación	53
3.5	Valoración del diseño de la evaluación	54
3.6	Los indicadores: instrumentos para la evaluación	55
3.6.1	Indicadores sociales	56
3.6.2	Atributos de los indicadores	57
3.7	Síntesis capitular	58
	Referencias	58
CAPÍTULO 4.	Ética y evaluación	60
	Introducción	60
4.1	Generalidades acerca de la ética	61
4.2	El rol del evaluador	62
4.2.1	Experto	62
4.2.2	Educador	63
4.2.3	Representante	64
4.3	Las opciones y los nodos éticos de la evaluación	65
4.4	La difícil relación entre la ética, la política y las políticas	67
4.5	Garantizando la aplicación de los imperativos	68
4.6	Síntesis capitular	70
	Referencias	71
CAPÍTULO 5.	Evaluación basada en la teoría	72
	Introducción	72
5.1	La supremacía del diseño en el proceso de las políticas	73
5.2	La política pública como teoría	74
5.3	Un breve paréntesis acerca de la teoría y la causalidad	75
5.4	La evaluación basada en la teoría	77
5.4.1	Evaluación lógica	78
5.4.2	Evaluación teórica	78
5.4.3	Teoría del cambio	79
5.4.4	Evaluación basada en la teoría (EBT)	81
5.4.5	Análisis de la contribución	82
5.5.	Instrumentos que soportan la evaluación teórica/lógica	83
5.5.1	El modelo lógico	83
5.5.2	La evaluación teórica de Weiss	84
5.5.3	El árbol de problemas	86

5.6	Importancia de evaluar los supuestos del diseño	88
5.7	Síntesis capitular	89
	Referencias	89
CAPÍTULO 6.	La evaluación de la efectividad	91
	Introducción	91
6.1	Significado de la efectividad	92
6.2	Antecedentes de la efectividad	93
6.3	Determinantes de la efectividad	95
6.4	La gerencia del desempeño	98
6.4.1	Técnicas para investigar el desempeño	100
6.4.2	Desempeño basado en evidencias	101
6.4.3	Retroalimentación 360°	101
6.4.4	<i>Coaching</i>	102
6.5	Medidas del desempeño	103
6.5.1	Las metas organizacionales	104
6.5.2	Tipos de metas	105
6.6	Síntesis capitular	105
	Referencias	106
CAPÍTULO 7.	Evaluación de la eficiencia económica	108
	Introducción	108
7.1	La eficiencia económica	108
7.2	Conceptos básicos de ingeniería económica	110
7.2.1	Principios del dinero	111
7.2.2	Valor presente y valor futuro	113
7.2.3	Serie uniforme	114
7.3	Análisis costo–beneficio	117
7.3.1	Valor Presente Neto (VPN)	118
7.3.2	Relación beneficio–costo (B/C)	120
7.3.3	Tasa Interna de Retorno (TIR)	121
7.3.4	Indicador de costo–eficiencia (C/E)	124
7.4	Evaluación económica en el sector público	127
7.5	Síntesis capitular	128
	Referencias	128
CAPÍTULO 8.	Evaluación económica del programa de la tortilla	130
	Introducción	130
8.1	Lineamientos del programa	130

8.2	Presupuesto del programa	132
8.3	Inversiones	134
8.4	Costos de operación	134
8.5	Beneficios	136
8.5.1	Familias atendidas	136
8.5.2	Kilogramos de tortilla abastecidos	137
8.5.3	Subsidio otorgado	137
8.6	Indicadores de evaluación	138
8.7	Evaluación costo-eficiencia del programa de la tortilla	139
8.7.1	Indicador de costo-eficiencia por familia atendida (C/E_i)	139
8.7.2	Indicador de costo-eficiencia por kilo entregado	140
8.7.3	Indicador de costo-eficiencia por peso subsidiado	140
8.7.4	Relación entre los indicadores de costo-eficiencia	141
8.8	Evaluación global del programa	143
8.9	Síntesis capitular	143
	Referencias	144
CAPÍTULO 9.	Evaluación política	145
	Introducción	145
9.1	Las políticas públicas como discurso	146
9.2	El discurso como foco evaluativo	149
9.3	Las dimensiones políticas de la evaluación	151
9.3.1	Democracia y políticas públicas	152
9.3.2	Deliberación y evaluación	152
9.3.3	Participación y evaluación	154
9.3.4	Transparencia y evaluación	155
9.3.5	Rendición de cuentas y evaluación	156
9.3.6	Buena gobernanza y evaluación	156
9.4	La democracia y sus indicadores	157
9.4.1	La medición de la democracia	158
9.5	Síntesis capitular	160
	Referencias	161
CAPÍTULO 10.	Evaluación de impacto	163
	Introducción	163
10.1	La medición del impacto	164
10.2	Proceso de evaluación del impacto	165
10.3	El asunto de los contrafactuales	167

10.4	Técnicas para la evaluación de impacto	169
10.4.1	El método de la diferencia en las diferencias	170
10.4.2	Método de regresión discontinua	173
10.5	Aplicación práctica	177
10.6	Síntesis capitular	186
	Referencias	187

JESÚS CARRILLO RODRÍGUEZ,
in memoriam

Presentación

El desempeño de las organizaciones sociales, ya sean privadas o públicas, está en permanente vigilancia de propios y extraños. En efecto, a diario, sus *stakeholders* exigen de su actuación mejores resultados, eficiencia creciente e impactos incuestionables. En ese sentido, quienes lideran dichas organizaciones enfrentan desafíos crecientes para alcanzar el reconocimiento y sentido necesarios de su gestión. Nada más necesario que la evaluación permanente para lograr que sus valores institucionales, su misión y su visión, estén soportados en procedimientos rigurosos, herramientas adecuadas y acciones permanentes que les faciliten su labor. La cultura de la evaluación requiere de una gran cruzada de afianzamiento para enfrentar, de manera adecuada, las exigencias del mundo actual. Hacer de la evaluación un hábito y una manera de pensar y de actuar no es nada elemental.

Sin duda, evaluar es una de las principales acciones del género humano y, en efecto, los seres humanos permanentemente enfrentan juicios sobre distintas esferas de sus vidas y sus contextos. Así, por ejemplo, se evalúa el sitio de estudio, de trabajo o la ciudad en la que viven. Asimismo, se juzga el comportamiento de otros o se determina cómo salieron los planes de distinta índole. Por supuesto que, tanto los gobiernos como sus decisiones y actuaciones, estarán bajo el escrutinio de los ciudadanos y no siempre con las mejores calificaciones. Indistintamente de los roles y condiciones, las personas emiten opiniones en ese intento de calificar los eventos acaecidos. En fin, la evaluación es una práctica habitual que se realiza con diversos instrumentos, sofisticados o no, y con distintos propósitos, unos buenos y otros no tanto.

Esta obra versa acerca de la evaluación y, en especial, de la valoración de la acción gubernamental. Intenta también ser un libro de consulta para quienes están interesados en estudiar el desempeño de las agencias gubernamentales, pero desde la óptica de las políticas públicas. Asume, como supuesto central, que la función evaluativa es multidimensional y que, en consonancia, exige y posibilita diversas formas de análisis y juicio. En efecto, para emprender un proceso de evaluación deben considerarse varias posibilidades, escenarios y criterios, a fin de emitir conceptos sobre la efectividad, la eficiencia, el impacto y, en general, sobre el buen gobierno. En los tiempos actuales, el buen gobierno es una exigencia generalizada, así como in-

terés supremo y preocupación creciente de los ciudadanos. De igual manera, el escrutinio del desempeño gubernamental en Latinoamérica ha hecho que la evaluación de las políticas asuma un papel protagónico en la opinión pública y que dicha valoración sea reconocida como una de las más promisorias estrategias para avanzar y mejorar. La conexión entre evaluación y desarrollo, en sentido amplio, está establecida de manera clara.

Esta obra refleja un esfuerzo de más de 30 años de dedicación y experiencia en el estudio del desempeño de las organizaciones privadas y públicas. Fue, sin duda, la formación como ingeniero industrial la que señaló la evaluación como un asunto de la mayor importancia y fue también su dimensión económica la que, en sus inicios, atrajo la atención. Después, tal interés se mantuvo durante la inserción en las agencias gubernamentales colombianas desde el Plan Nacional de Rehabilitación (PNR) y, en especial, cuando desde el Banco de Proyectos de Inversión Nacional (BPIN) del Departamento Nacional de Planeación (DNP) se propuso la cultura de proyectos como una estrategia para mejorar el uso de los recursos públicos. Los estudios superiores reforzaron el compromiso con las actividades evaluativas y animaron la incursión de manera decidida como docente e investigador en dicho campo.

Desde entonces, el compromiso con la función evaluativa ha sido compartido en los diversos cursos orientados en las distintas facultades mexicanas y colombianas que brindaron tal oportunidad. Como docente, he pretendido acrecentar la masa crítica de evaluadores, reclutar los mejores aliados para socializar, inducir y conseguir óptimos resultados en las intervenciones públicas. Estoy convencido de que hacer un uso adecuado de los recursos, mejorar permanentemente los procesos y garantizar efectividad, son derroteros que debieran perseguirse a diario. En esencia, la evaluación promueve la consecución de tales aspiraciones. En sus técnicas y perspectivas subyace un afán por generar la información para procurar decisiones orientadas al cumplimiento adecuado de competencias y mandatos constitucionales. Aun así, hay que decirlo, no todos sienten simpatía por la evaluación, sobre todo, los gobernantes.

El contenido del texto propone una visión integral del desempeño institucional. Por ello, intenta dar respuesta a sendos interrogantes propuestos para el desarrollo de las distintas temáticas. Se pregunta, en su inicio, qué es la evaluación, cómo ha evolucionado y desde qué campos disciplinarios se promueve su dinámica. Posteriormente, da respuesta a cómo se clasifica y qué métodos son aplicados; enseguida, responde a cómo se diseña y cuáles son los principales componentes para emprender un ejercicio evaluativo. Después, se dedica a las diversas técnicas para emitir juicios acerca de los distintos atributos de una política pública. Para responder la pregunta de cómo evaluar, es conveniente advertir que se asume que las técnicas elegidas están asociadas al concepto de política pública. Como se sabe, no hay consenso sobre tal concepto; ello da cuenta de lo nueva que es la disciplina, pero también de los distintos significados que pueden adoptarse.

Es importante advertir que para el desarrollo del contenido y la aplicación de las distintas técnicas de evaluación se parte de la premisa que el significado de política pública condi-

ciona su aplicación. Así, por ejemplo, si se asume la política como una competencia o responsabilidad de una agencia gubernamental, se estará evaluando el desempeño y su capacidad organizacional para dar cumplimiento a dichas responsabilidades. En otras palabras, se evaluará la efectividad. Si, por otra parte, se asume una política como producto de una teoría del cambio social, se evaluará su coherencia y fortaleza teórica. Por otro lado, si la política es un flujo –entradas y salidas– de recursos, deberá evaluarse también la eficiencia del uso de tales recursos. Al considerar la política como un discurso político, se juzgarán los valores que persigue o pretende el arreglo discursivo. En resumen, esta perspectiva integral da cuenta de las variadas posibilidades de la evaluación y de las distintas técnicas a disposición. Para mayor ilustración, el último capítulo presenta algunas experiencias reales de la evaluación.

Esta publicación está en deuda con muchas personas e instituciones, pues sin ellas no hubiera sido posible que esta saliera a la luz. En primera instancia, con aquellos estudiantes de las diversas licenciaturas, maestrías y doctorados, quienes participaron en los cursos impartidos en las distintas universidades de Colombia y México. A ellos, mi gratitud por su atención, interés y aportes a las diversas temáticas discutidas, así como por difundir, en sus propios contextos, la importancia del ejercicio evaluativo. Es pertinente hacer mención de las instituciones educativas que han tenido la generosidad de acogerme y fomentar mi actividad docente: Universidad de Cundinamarca, Pontificia Universidad Javeriana, Universidad Sergio Arboleda, la Escuela Superior de Administración Pública (ESAP) y la Universidad Externado de Colombia, todas en Colombia. En México, la Facultad Latinoamericana de Ciencias Sociales (FLACSO), la Universidad Autónoma de la Ciudad de México (UACM), la Universidad Tecnológica de México (UNITEC), el Instituto Nacional de Administración Pública (INAP), la Universidad Anáhuac México y la Universidad Iberoamericana Puebla. A mis colegas de todas ellas, mi gratitud y aprecio por sus enseñanzas y apoyo permanente.

CAPÍTULO 1

Conceptos y antecedentes

INTRODUCCIÓN

La evaluación de políticas (*policy evaluation*) vive uno de sus mejores momentos en toda su historia. Su ascenso como campo disciplinar puede constatarse, entre otras cosas, por la producción intelectual especializada, el notable número de organizaciones que fomentan su aplicación y su adopción formal en los distintos niveles y agencias gubernamentales. Este capítulo presenta las premisas básicas del significado de la evaluación, su recorrido histórico y los ámbitos donde se desenvuelve; bien sea como campo disciplinar, profesión o función organizacional. También presenta un breve recorrido de la evaluación por algunos campos disciplinares, intentando resaltar la no exclusividad de sus dominios con una única disciplina. Como todo capítulo introductorio, este intenta servir como basamento para el desarrollo posterior de algunos elementos más particulares y densos de la actividad evaluativa.

1.1 CONCEPTOS BÁSICOS

A fin de establecer las claridades conceptuales necesarias y convenientes, se propone una reflexión del significado de la evaluación en el ámbito cotidiano y en los distintos campos disciplinares que reconocen la función evaluativa. Esto, para ratificar que la actividad tiene una gran ascendencia en la vida de los individuos, las organizaciones y los gobiernos. Asimismo, es una actividad permanente, con distintas modalidades y propósitos que, en últimas, conduce a decisiones ulteriores y ocasiona impactos significativos para las actividades que realizan los individuos. Entonces, no puede considerarse una actividad neutral, ya que produce efectos e impactos para quienes son sus sujetos, y también puede cambiar cursos de acción, relaciones y destinos. Se evalúa con dicha finalidad y debemos atenernos a sus consecuencias.

En el nivel más básico, “evaluar” significa “valorar”. De manera amplia, dicha valoración puede darse en los campos del conocimiento, aptitudes, capacidades o rendimientos de cualquier actividad o emprendimiento. Evaluar también significa juzgar qué tan buena, útil o

exitosa resulta una actividad. La evaluación, en consecuencia, es un procedimiento que permite la generación de juicios acerca de una situación específica. Tal y como fue expresado, puede estar vinculada a diferentes esferas de la actividad humana o a los aspectos del contexto donde se desarrolla. Por lo regular, en el ámbito académico se pueden encontrar desde las conceptualizaciones más elementales hasta las más sofisticadas, aunque en ellas permanezcan los elementos esenciales o los propósitos buscados.

La naturaleza polisémica del término “evaluación” es una de sus principales características y alude a muchas otras acepciones del contexto o del campo en el que se refiere. Es así como por evaluación puede entenderse una estimación, un cálculo o un juicio de valor sobre algún atributo o propiedad de algo o de alguien. De igual modo, a dicho término se asocian el juzgar o determinar el significado, así como su valor o calidad. La condición de quien expresa el juicio o valoración también cuenta; en consecuencia, se insiste en el carácter oficial del asunto, muy seguro haciendo referencia a la autoridad, idoneidad o importancia de quien los emite.

La evaluación refiere una actividad funcional que puede realizar cualquier persona en cualquier ámbito de la vida. Así, por ejemplo, un evento o suceso se juzga o valora en importancia o mérito. En todo caso, es el acto o la instancia de evaluar, emitir juicio o diagnóstico de una condición física o mental. No puede desconocerse que este ejercicio podría tener una gran dosis de discrecionalidad o ser producto de cargas ideológicas, en la medida que se base en evidencias anecdóticas o fragmentarias o en criterios particulares y ligeros. Sin duda, en ocasiones puede cuestionarse la objetividad de la evaluación.

Para la acción pública, la evaluación significa mucho más y constituye un gran desafío para gobernantes y gobernados, y en torno a la que surgen disputas sociales. Desde estos escenarios, la evaluación es un ejercicio selectivo que intenta, sistemática y objetivamente, estimar el progreso hacia algún fin determinado y el alcance de dichos resultados. De ninguna manera puede considerarse unidimensional, por el contrario, involucra miradas distintas sobre diferentes espectros y con diversos niveles de profundidad sobre los atributos conceptuales. La sincronía no es una condición necesaria, ya que pueden valorarse en distintos puntos de tiempo o atributos. Su carácter dinámico también es una cualidad notoria; revelando, por tanto, que las cosas pueden cambiar, evolucionar o adaptarse, condiciones propias de la vida en sociedad y de la acción política. No hay nada más dinámico que el juicio sobre el desempeño político.

1.2 ORÍGENES DE LA EVALUACIÓN

La creciente importancia de la evaluación no significa que su estatuto disciplinar sea reciente, ya que presenta importantes antecedentes en la historia de la humanidad. Al parecer, es tan antigua como el mismo ser humano, pues todas las actividades realizadas son susceptibles

de evaluación y siempre hay inclinación hacia la presentación de balances al final de la jornada, cualquiera que sea. Muchas disciplinas como la estadística han fomentado de manera notoria la evaluación. En paralelo, dadas las actividades profesionalizantes e investigativas de las ciencias sociales, hay “mayor disponibilidad de técnicas y campos de acción que reclaman la función evaluativa” (Carrillo, 2009, p. 56). Redondeando, “se tienen antecedentes significativos que pueden dar cuenta de su autonomía como campo disciplinar, como los que registra Carol H. Weiss (1998), quien ubica la prehistoria de la evaluación hacia 1660” (Carrillo, 2009, p. 56). La destacada investigadora propone el inicio con “los estudios empíricos de los problemas sociales en Gran Bretaña y con el intento de conocer las relaciones entre sociedad y economía” (Carrillo, 2009, p. 56).

John Graunt (1620-1674) y William Petty (1623-1687) son reconocidos como los baluartes de la estadística científica a partir de su análisis de los boletines de mortalidad (*Bills of mortality*), cuya publicación en 1662 se hizo bajo el nombre de *Natural and Political Observations*. En ese momento se reportaron las principales causas de muerte durante más de 50 años. Con estas obras se inaugura un campo interesante para la investigación y la actividad gubernamental, dándose a conocer con el nombre de aritmética política (*Political Arithmetic*), para denominar aquellas actividades oficiales emprendidas, a fin de generar información relevante sobre la marcha económica de una nación.

Un trabajo realmente evaluativo, según Weiss (1998), es el presentado por André-Michel Guerry, en 1833, sobre lo que denominó la estadística moral de Francia, en él se presentaron asuntos relacionados con fenómenos como el crimen, los suicidios y el analfabetismo, entre otros. Este trabajo desarrollado entre 1820 y 1830 utilizó mapas y tablas para analizar una variedad de problemas sociales de manera comprehensiva. En la actualidad, entre los especialistas se reconoce como el estudio que inauguró las ciencias sociales empíricas modernas (Friendly, 2007). En general, el estudio de Guerry constituye una contribución importante a la aritmética política. Bajo esta disciplina, los gobernantes deberían investigar, cuantificar y registrar todos los recursos a su disposición (población, tierras, tributos, transacciones comerciales, etcétera).

Otro famoso estudio de carácter evaluativo fue el liderado por Jules Deputit (1804-1866), connotado ingeniero y economista francés que estudió la utilidad del trabajo público mediante el cálculo económico. Esto, al punto de tener un lugar destacado en la historia del pensamiento económico y ser considerado el precursor del análisis marginal. En su estudio, *On the Measurement of the Utility of Public Work*, se propuso demostrar las formas de medir la utilidad de un puente peatonal. Para ello supuso que la utilidad es un atributo subjetivo; esto es, varía de acuerdo con los usuarios, vinculándolo a la disposición para pagar por un determinado bien. Dichos estudios presentaron “el uso de técnicas estadísticas, pero en verdad, no pretendían un ejercicio evaluativo y distaban mucho de lo que hoy en día se reconoce como tal en algunas disciplinas de las ciencias sociales” (Carrillo, 2009, p. 57). Un análisis detallado y juicioso de la obra de Deputit puede apreciarse en Ekelund (1968).

1.3 LA EVALUACIÓN Y SUS DOMINIOS

La evaluación no es una función exclusiva de una actividad humana o disciplina académica en particular; por el contrario, está presente en muchas de ellas, aunque de manera inconsciente. Es tan común que puede considerarse permanente, ya que el ser humano en ejercicio de su cotidianeidad evalúa todo el tiempo su sitio de residencia o de trabajo, sus decisiones de compra, venta, etc. Al hacerlo está examinando y valorando un fenómeno a la luz de algún patrón o norma, explícito o implícito, que puede involucrar categorías como: estética, ¿es bueno y placentero?; efectividad, ¿lo que está haciendo es lo que se supone debería hacer?; eficiencia, ¿proporciona los beneficios de acuerdo con los costos?, o también otras como justicia, equidad, aceptabilidad, etc. Por tanto, una actividad de evaluación subyace a cualquier ejercicio de toma de decisiones. En realidad, no puede afirmarse que la evaluación tenga un origen disciplinar único; por el contrario, durante su expansión y consolidación puede constatar que distintas disciplinas han acogido los preceptos evaluativos, incorporado sus propias visiones y, en ocasiones, adecuándolos a sus fines. A continuación, se presenta una breve reseña de este tránsito por los campos disciplinares relacionados con el ejercicio.

1.3.1 Ciencias administrativas

Quizás uno de los vínculos más fuertes de la evaluación es el que existe con las ciencias administrativas. Es claro que ambas están en estrecha relación con la historia de las actividades humanas y son palpables en su recorrido histórico. El trabajo y la supervivencia motivaron al ser humano a encontrar formas creativas y sistemáticas para desarrollar sus actividades; indistintamente enfrentarán tiempos de guerra o la simple cotidianeidad. Por ello, se reconoce que la administración y la evaluación han estado vinculadas desde siempre al ser humano y han sido consustanciales a su evolución.

Quienes practican la arqueología administrativa dan cuenta de actividades de planeación, dirección y control en importantes emprendimientos humanos como las pirámides egipcias, el auge y la expansión de los imperios, así como en la función pública de los incipientes Estados. Muchos periodos históricos señalados en la humanidad soportan su división en los mayores avances, claramente relacionados con la evolución de la administración. Por ejemplo, la Revolución Industrial es reconocida por importantes avances en la economía y la industria, basados estos en aportes científicos y en la manera de organizar las actividades laborales. Después, las ciencias administrativas son las que incorporan la evaluación al proceso gerencial/administrativo, conformado por cuatro funciones principales: planeación, organización, dirección y control, siendo esta última la que incluye las funciones evaluativas.

En la historia del pensamiento administrativo, Henry Fayol (1841-1925) y Frederick W. Taylor (1856-1915) son reconocidos por sus aportes a la teoría de la administración científica. Ambos ingenieros estudiaron la industria de la minería y observaron el ambiente labo-

ral desde distintas latitudes, sentando las bases de la administración científica. La obra de Fayol, “Administración industrial y general”, publicada en Francia en 1916, desarrolló un modelo administrativo basado en cinco funciones de la gerencia: planeación, organización, integración, dirección y control. La perspectiva de los autores proponía la generación de una revolución mental y de un trabajador funcional a partir de los principios de la administración, generando un nuevo modelo industrial que privilegió la expansión y el crecimiento de las grandes potencias. Desde entonces, la función de evaluación, en este caso del desempeño, ha ocupado un lugar preeminente en las cuestiones administrativas.

En el periodo de “entre guerras”, varios estudios estuvieron influidos por el enfoque evaluativo, destacando entre ellos los realizados en el campo de la experimentación social y emergidos desde la psicología, así como los efectuados en materia de productividad laboral en la Western Electric (Hawthorne Effect). Apartado de la visión clásica de la administración (énfasis en la tecnología y los procesos como determinantes de la productividad), el enfoque de las relaciones humanas, también reconocido como el “enfoque humanista de la administración”, propuso principios del desempeño laboral a partir de un diseño experimental. Elton Mayo (1880-1949) y su equipo de colaboradores emprendieron un experimento de largo aliento que, al correlacionar la productividad y algunos factores ambientales (iluminación, ventilación, sistemas de remuneración, periodos de descanso, etc.) en dos grupos de trabajo, crearon un índice de comparación. Mas adelante, se dedicaron a estudiar la influencia de los grupos en el comportamiento individual, así como las actitudes y los sentimientos de los trabajadores, concluyendo en la enorme influencia del comportamiento grupal sobre el individual. Desde entonces, la teoría administrativa ha considerado de manera creciente estas influencias y ha postulado que los administradores no pueden tratar a los trabajadores de manera aislada, como si fueran átomos, pues ellos se deben a un grupo y a una sociedad, que de alguna manera impone sus parámetros.

A mediados del siglo xx toma lugar el *boom* de la evaluación, sobre todo en el campo social. Durante dicho periodo, fundaciones privadas empezaron a evaluar los fondos disponibles; a la par, el gobierno federal de Estados Unidos patrocinaba nuevos esfuerzos evaluativos. En los años sesenta, la publicación de libros y artículos acerca de dicho tema repuntó, de modo que la RAND Corporation y el Departamento de Defensa, entre otros, adelantaron análisis costo-beneficio. Poco después, se institucionalizó la evaluación en las actividades gubernamentales y surgieron las asociaciones de evaluadores. Sin duda, el avance de este campo de estudio fue posible gracias a los métodos de investigación y a la aplicación de la estadística para el estudio de los problemas y los procesos sociales y las relaciones interpersonales. Es así como todo este desarrollo ha estado acompañado y ha sido estimulado por mejoras en las técnicas de recolección de datos y por el uso de las computadoras para el análisis de una gran cantidad de datos (análisis multivariado).

En la actualidad, en las ciencias administrativas (de negocios, empresas, organizaciones, etc.), la evaluación es una función esencial y parte neurálgica de los procesos de gestión.

Aunque muchos autores refieren el control en lugar de la evaluación, la naturaleza y propósitos entre las dos etapas del proceso permanecen difusos y sus fronteras no se determinan con facilidad. En la administración, el foco del “control” es la evaluación y corrección del desempeño de los subordinados para asegurar que los objetivos y planes se lleven a cabo (Koontz y O’Donnell, 1975). Aunque no constituye un propósito inicial dilucidar esta superposición de términos; de momento, basta señalar la vinculación íntima de los términos y el uso conceptual sin distingo. De igual modo, sería necesario resaltar que la ausencia de la “evaluación” en el proceso administrativo lo deja inconcluso y puede impedir el logro de las metas propuestas en materia de productividad. Al no desplegarse como una etapa, resulta imposible determinar el nivel de éxito y establecer los correctivos necesarios.

1.3.2 Desarrollo económico

La referencia del desarrollo económico es un requisito fundamental de la evaluación de programas sociales, políticas públicas y de todo tipo de acciones gubernamentales; son tan fuertes sus vínculos que pocos desconocerían la pertinencia de su abordaje. Su consideración es imperiosa si se quiere abordar integralmente la concepción de las iniciativas gubernamentales, en especial, su naturaleza extensa, las controversias en sus visiones y la extraordinaria dinámica que en ciertas ocasiones dificulta su seguimiento. En la actualidad se enfatiza la necesidad de potenciar capacidades evaluativas en las agencias gubernamentales y las de cooperación internacional, ya que la función ha propiciado la búsqueda de mayores niveles de desempeño en algunos países mediante el conocimiento de variables intervinientes, refinamiento de postulados y desarrollo de nuevos constructos. La evaluación ha sido de gran utilidad para la cooperación del desarrollo.

Durante las últimas décadas se han introducido nuevos requisitos para concebir el desarrollo y para lograr las condiciones que lo materializan. En América Latina, la preocupación ha sido mayor y existe evidencia desde las “misiones económicas” establecidas en los distintos países y las organizaciones a cargo de ellas, como la Comisión Económica para América Latina (CEPAL). Gracias al trabajo de estas comisiones se logró el consenso sobre las condiciones que promueven el desarrollo en cualquier sociedad. La estabilidad macroeconómica es requisito esencial para alcanzar el crecimiento necesario para el desarrollo. Contrario a lo que se postuló en épocas pasadas, el desarrollo no se mueve desde las capas más ricas hacia las más pobres, o desde los países ricos hacia los pobres, para finalmente tocar a todos los grupos en desventaja. Existe plena conciencia que el desarrollo se puede propiciar para resolver, de manera directa, las necesidades humanas y que no existe una política única que lo detone, por lo que se demanda un enfoque comprensivo. Debe concebirse como un proceso sostenible originado en prácticas socialmente inclusivas y sensibles para cambiar las tendencias. En este sentido, las instituciones determinan el desarrollo en gran medida.

La historia del desarrollo está fuertemente vinculada a las ciencias económicas, y desde allí se ha animado su debate. Tras salir del feudalismo, las nuevas condiciones indicaron la pertinencia de explicar el crecimiento económico a partir de los factores de producción y del precio de los bienes. Esta época corresponde a los desarrollos teóricos de Smith, Ricardo y otros notables pensadores de la *economía clásica*. En estos primeros esbozos, la inquietud fue evaluar el desempeño de las economías como un todo, estudiar el uso de los recursos disponibles y determinar la riqueza de las naciones. Posterior a ello, la denominada *economía neoclásica* fundamentó el conocimiento sobre el comportamiento económico, basado en esencia en la conducta del individuo. Bajo los enfoques marginalistas de la utilidad se asumió la competencia perfecta y otros supuestos para profundizar en torno al desarrollo, al uso de los recursos y a las necesidades de la población. Estos adelantos permitieron reconocer el papel de las instituciones y las motivaciones subyacentes en la vida social y de los gobiernos en el bienestar social.

A pesar de sus debates e imperfecciones, la teoría económica ha servido para entender, en buena parte, los problemas de vivir en sociedad. Una vez terminada la Segunda Guerra Mundial, es cuando los temas del desarrollo surgen con ímpetu desmesurado. Con un nuevo panorama geopolítico y el plan Marshall (1948) para la reconstrucción de Europa occidental y para pensar en la cooperación internacional, los temas del desarrollo enfatizaron la diferencia entre países pobres y ricos, proponiendo así la ayuda externa a los primeros para escapar de la trampa de la pobreza.

Como zona particular, América Latina ha estado en los radares de las agencias multilaterales y sus procesos económicos han dependido, en buena manera, de la incidencia externa. Durante la industrialización dirigida por el Estado, la economía latinoamericana siguió creciendo algo más que el promedio y su participación en la producción mundial alcanzó 9.5% en 1980 (Ocampo, 2012). Es desafortunado, pero la historia constata que los resultados no fueron tan halagadores como se esperaba, ya que los países latinoamericanos estaban tan empobrecidos que fueron incapaces de alcanzar las metas de crecimiento y ahorro, y los resultados de la estrategia fueron lamentables. Sin embargo, estas experiencias sirvieron para refinar la teoría, al reconocer otras variables políticas, sociales y éticas que limitaban y obstaculizaban el avance y desarrollo en los países pobres.

La teoría de la dependencia para comprender e interpretar el desarrollo latinoamericano fue con base en otra perspectiva; en especial, cuando las medidas sugeridas fracasaron. Las propuestas de Gunder, Prebisch, Dos Santos, Cardoso y Faletto, entre otros, son interpretaciones para entender las relaciones centro-periferia (CLACSO, 2017). Para elevar los niveles de la calidad de vida, los trabajos de los mencionados autores combinaron elementos neomarxistas con la teoría keynesiana para fortalecer la demanda interna, desarrollar mercados industriales, mejorar la renta de los trabajadores y elevar la efectividad gubernamental. La teoría de la dependencia, entonces, evaluaba las relaciones entre países desarrollados y los menos desarrollados, constituyéndose en una alternativa a la teoría de la modernización preconizada

desde los países industriales. En ella prevaleció la visión histórica sobre el desarrollo económico, destacando los roles del centro y la periferia, estos últimos como proveedores de materias primas y receptores de productos terminados una vez saturados los mercados del norte. En resumen, estos países “dependen” de la voluntad de compra de los países industrializados.

Aunque la acumulación del capital físico y del humano es esencial para el crecimiento económico, su presencia para la medición del desarrollo ha cambiado. La evidencia empírica muestra que países con altas tasas de acumulación físico tienen un crecimiento económico modesto. De igual modo, existen casos donde el crecimiento no ha tomado lugar, a pesar de mejoras y grandes esfuerzos educativos. A la luz de estas paradojas se incorporan nuevas visiones acerca del desarrollo, mismas que dieron cuenta de la estructura de incentivos que conduce a la acumulación de capital físico y humano, y al progreso tecnológico. Bajo esta mirada, la atención se ha centrado en el análisis de las políticas económicas, las instituciones políticas y económicas y el capital social (Coleman, 1988). De esta forma, la ampliación teórica del desarrollo ha incorporado cuestiones del desarrollo sostenible y de una nueva racionalidad ética de la economía, incorporando la justicia social y económica y el desarrollo humano.

[L]as crisis económicas y políticas en los países [de América Latina] han [desembocado en] medidas tendientes al logro de un Estado con niveles aceptables de eficiencia y efectividad. La discusión sobre la transformación del aparato estatal ha transitado por temas relacionados con el tamaño, con sus competencias y con las medidas necesarias para conjurar los vicios incubados en la administración pública (Carrillo, 2009, p. 54).

A pesar de su reciente introducción, el despliegue de la evaluación en el sector público tiene hondas repercusiones en los distintos niveles de la gestión gubernamental y promete interesantes incidencias en el desempeño de las agencias estatales. Se espera tanto un proceso de aprendizaje continuo que retroalimente la etapa de planeación, como una generación de información como insumo para la toma de decisiones que apunten al mejoramiento continuo de la gestión pública.

1.3.3 Educación

La evaluación en el campo educativo es una práctica regular y después de más de 200 años de aplicación puede considerarse una actividad madura con importantes adelantos y perspectivas. En el mundo anglosajón, la evaluación de los programas educativos atraviesa siete periodos de desarrollo (Madaus *et al.*, 2000 en Hogan, 2007, p. 3).

1. *Age of reform* (anterior a 1900). Fue propiciada para medir el desempeño estudiantil y para fomentar la sicometría en Inglaterra. En Estados Unidos, la primera eva-

luación tomó lugar en Boston, mediante un test impreso en 1845, inaugurando así una larga tradición de la evaluación del aprendizaje.

2. *Age of efficiency* (de 1900 a 1930). En este lapso se tomó como referencia el trabajo de W. Taylor, el cual se aplicó a los estudiantes de administración. Para entonces, medición y evaluación eran sinónimos.
3. *Tylerian age* (de 1930 a 1945). En este periodo, los objetivos instruccionales creados por Ralph Tyler se tomaron para la evaluación educativa.
4. *Age of innocence* (de 1946 a 1957). Se caracterizó por la expansión de la oferta educativa y del gasto público en el sector. Asimismo, se consolidó el Tylerismo y se adicionaron propuestas como la taxonomía de los objetivos educacionales de Bloom, Engelhart, Furst, Hill y Krathwohl (1956).
5. *Age of development* (de 1958 a 1972). En estos años se dio un fuerte apoyo a las áreas consideradas estratégicas para la seguridad nacional; por tanto, la evaluación fue soportada con fondos públicos.
6. *Age of professionalization* (de 1973 a 1983). En este periodo, la evaluación emergió como una profesión y los *journals* especializados empezaron a tener auge y aparecieron los primeros programas universitarios para formar evaluadores educativos.
7. *Age of expansion and integration* (de 1983 a 2000). Se conformaron las primeras asociaciones de profesionales y se crearon comités conjuntos para la evaluación del personal. Durante esta época, la evaluación se extendió e integró a otras dimensiones de la actividad humana.

Para algunos autores, la actividad de evaluación en la educación encarna ciertos riesgos que deben atenderse (Worthen, Sanders y Fitzpatrick, 1997 en Hogan, 2007). Esto indica que la actividad ha ganado en legitimidad, por lo que en algunos contextos es reconocida sin mayores problemas, al complementarse con alternativas cualitativas integradas a las tradicionales cuantitativas (métodos multivariados). Como actividad profesional tiene un soporte teórico importante acompañado de estándares éticos que la regulan.

1.3.4 Investigación

La importancia de la evaluación es tal, que ha configurado un nuevo campo disciplinar: la investigación evaluativa. Asume, por tanto, un lugar privilegiado en el análisis de las políticas y programas gubernamentales y se ocupa de recolectar, verificar e interpretar la información en torno a su ejecución y eficacia. Existe una diferencia entre emitir juicios de valor sin considerar su base y hacer investigación evaluativa, en la medida que se apoya en métodos científicos para propósitos específicos (Suchman, 1967 en Joyce, 1980). Esto, precisamente

es lo que intenta superar la investigación evaluativa, que los juicios tengan una base científica y se alejen de especulaciones y sesgos innecesarios.

Es muy común encontrar en la bibliografía especializada el concepto integrado de investigación evaluativa (*Evaluation Research*), debido a que, por separado, “investigación” y “evaluación” comparten algunas particularidades. Utilizan técnicas cuantitativas y cualitativas, independientemente del propósito que persigan, son rigurosos en su esencia y tratan los fenómenos de manera sistemática. Asimismo, emplean métodos para recolectar información, procesar y analizar datos y presentar resultados en esquemas más o menos similares. La evaluación como un tipo de investigación intenta describir una situación, establecer relaciones entre variables y, en ocasiones, trazar secuencias causales entre variables. Como diferencia, se encuentra que la evaluación se usa o aplica, mientras que la investigación genera conocimiento para después diseminarlo y aplicarlo.

La evaluación investigativa se entiende, entonces, como “la aplicación sistemática de las ciencias sociales y de sus procedimientos para evaluar la concepción y el diseño, la implementación y el impacto social de las intervenciones públicas” (Davis, 1990, p. 35). En esta misma dirección, la evaluación de programas se define como “el uso de los procedimientos de la investigación social para indagar sistemáticamente sobre la efectividad de sus intervenciones” (Rossi, 2008, p. 4). En su forma más elemental, la evaluación está hecha para un cliente que ha decidido emprenderla, por lo que la considera como un medio para generar información y soportar sus decisiones.

1.3.5 Políticas públicas

Dada la naturaleza de la actividad pública, la evaluación de políticas es tan antigua como los gobiernos mismos. Al fin y al cabo, legisladores, personal administrativo, jueces, grupos de presión, medios de comunicación y ciudadanos siempre han esgrimido juicios acerca del valor o los efectos de políticas públicas, programas o proyectos particulares (Anderson, 2006). La evaluación tiene tanto de arte como de ciencia y comprende la estimación, juzgamiento o valoración de la política, de su contenido, implementación y del alcance de las metas y otros efectos (Anderson, 2006).

La “ciencia de las políticas”, como en un inicio se denominaron, datan de 1951, año en el que Harold Dwight Lasswell propuso un campo disciplinar dedicado a estudiar la acción gubernamental. La orientación hacia las políticas (así designada por Lasswell) tenía diferentes pretensiones en su desarrollo. Primero, intentaba dar respuesta a una crisis de seguridad nacional que indica la necesidad de usar con mayor eficiencia los recursos, usarlos con la “más sabia economía” (Lasswell, 1951, p. 71). Más allá de este afán político-económico, en segunda instancia debía atender preocupaciones propias de la academia, a saber: ¿Qué investigar? ¿Qué método(s) utilizar? ¿Cómo hacer que los resultados de la investigación incidan en las decisiones públicas? Subyacía, en todo caso, la intención de inducir racionalidad en el

proceso de toma de decisiones políticas y, por tanto, había que ofrecer un espacio para lo que denominaba la “función de inteligencia”, alejando así las decisiones políticas de todo capricho y visceralidad. La ciencia de las políticas de la democracia, como también denominó a sus propuestas, postulaba propósitos ulteriores para generar el conocimiento necesario para mejorar la práctica de la democracia, cuya meta final fuera la realización o el alcance de la dignidad humana, en la teoría y en los hechos (Lasswell, 1951).

En toda la agenda investigativa, el conocimiento era un ingrediente esencial para alcanzar los objetivos propuestos, pero su generación le tomaría a Lasswell un par de décadas más. Exigía entonces, no solo incorporar a las decisiones conocimiento, sino también, conocer el proceso de la toma de decisiones en el orden público y civil (Lasswell, 1971). El proceso estándar de las políticas públicas o sus niveles heurísticos se configura por una secuencia de etapas para reconocer cómo se trata un problema de índole pública de principio a fin y cómo las agencias pueden generar todo un proceso de decisión y uso de los recursos (Jones, 1969; De León, 1982; Nakamura, 1987 y Anderson, 2006). Desde esta perspectiva puede afirmarse que la visión adoptada es exclusiva de los gobiernos y las autoridades con facultades para dar tratamiento a los problemas de vivir en sociedad. Lo anterior es tan solo una perspectiva en la que los gobiernos tienen la exclusividad del proceso de las políticas públicas. En 1960, Lasswell propuso un conjunto de fases que iniciaban con la prescripción, la inteligencia, la recomendación, la invocación, la aplicación, la evaluación y la terminación, para conocer el proceso de las políticas. Después, dichas etapas asumieron algunas denominaciones producto de las propuestas de otros estudiosos. Como estrategia analítica, los especialistas dividieron el proceso en una serie de etapas, fases, pasos o niveles heurísticos para encontrar respuestas, tomar decisiones, actuar públicamente o mejorar el desempeño de las organizaciones. En realidad, esta preocupación fue la que en últimas permitió que el mismo Lasswell y sus sucesores determinaran y consolidaran el campo de las políticas públicas. Fue hasta la década de los setenta cuando se refinaron algunas de sus propuestas.

En *El futuro de la ciencia política*, Lasswell (1963) propuso una relación entre el especialista de la ciencia política y el proceso de decisión que iniciaba con un flujo de información de alguien afectado. Después, un conjunto de actividades influía sobre el desenlace (promoción) para continuar con la articulación de normas (prescripción), la invocación, la aplicación, la valoración y la terminación. Por último, la evaluación se asocia a la mencionada valoración que caracterizaba la relación entre los objetos políticos, las estrategias y los resultados obtenidos. Esta propuesta no fue definitiva y con el tiempo evolucionó y tomó otras denominaciones, lo que a la postre sirvió para desarrollar el campo disciplinar de las políticas públicas. A pesar de los cambios en su denominación, la evaluación siempre permaneció como una de las etapas y, desde entonces, se mantiene dentro del proceso de las políticas.

En esta expansión de las políticas públicas como campo disciplinar, la evaluación de programas ha sido considerada de distinta manera y con diversos énfasis, “es el uso de procedimientos de investigación social para estudiar sistemáticamente la efectividad de los progra-

mas de intervención social, investigación que es adaptada a su ambiente político y organizacional, y diseñada para informar la manera de mejorar las condiciones sociales” (Rossi, 1998, p. 20). De la definición previa, es importante resaltar el uso indistinto o equivalente de los términos, políticas públicas y programas, algo extraño para quienes se aproximan a la disciplina. De igual manera, es importante destacar el vínculo o la cercanía de la evaluación y la investigación social, ya que el procedimiento evaluativo toma prestados algunos asuntos de la investigación social, imprimiéndole rigor y aspiraciones científicas para emitir juicios sobre las iniciativas públicas. El autor aclara que la evaluación juzga la efectividad de los programas por encima de otros atributos. Si bien, esta apreciación ha despertado inquietudes y debates públicos desde distintas orillas, no se puede desconocer su trascendencia para entender la acción gubernamental, ya que a los ciudadanos solo les interesa que los gobiernos cumplan lo prometido y que sus intervenciones mejoren sus condiciones.

La evaluación de programas está relacionada con la investigación aplicada, usada como parte de un proceso gerencial y conducida para ayudar a quienes toman decisiones administrativas sobre un programa o un servicio (Barker en Royse, 2016). La evaluación es una función primordial de quienes están al mando o dirigen los programas y cuya obligación es velar por su futuro. Ellos tienen una responsabilidad, aunque parcial, sobre el futuro de la política, al determinar si continúa o no y, sobre todo, en qué condiciones lo haría. Obviamente, la evaluación ofrece los elementos para este tipo de determinaciones, al tomar como base el desempeño y los logros alcanzados. La esencia de la evaluación radica en la estimación sistemática de la operación y/o resultados de un programa o política, comparada con un conjunto (algunas veces explícitos y otras no) de estándares, como medio para su mejoramiento (Weiss, 1998). También puede apreciarse un vínculo estrecho a los procesos de planificación, los cuales determinan los propósitos de una acción gubernamental o pública. En general, la evaluación se efectúa sobre asuntos predeterminados y, a la vez, constituye un acto consciente para vincular decisiones inter-temporales de lo que se propusieron algunos actores para un periodo, establecido en ocasiones por la duración de un gobierno. Pero no siempre están establecidos todos los propósitos y la improvisación supera la planificación y eso dificulta la acción evaluativa. Es necesario reconocer que muchas iniciativas sociales son resultado de incoherencias y, en ocasiones, del capricho del gobernante. En tales situaciones, desde luego que los estándares no son explícitos, por lo que el evaluador tiene que reconstruirlos, suponerlos e, incluso, inventarlos para adelantar un ejercicio de comparación.

Para resumir, la evaluación de programas (*program evaluation*) y la investigación comparan propósitos y técnicas para determinar si un programa público funciona y en qué magnitud lo hace (Morehouse, 1972). En este sentido, la evaluación determina qué funciona y qué no, establece de manera sistemática si un programa gubernamental logra sus objetivos. Como etapa, la evaluación puede ser simple o compleja, a corto o largo plazo, superficial o profunda, según sea la demanda. En esencia, proporciona evidencia sobre la efectividad de los programas para que los interesados reclamen su continuidad o mejora (Berner y Bronson, 2005).

1.4 SÍNTESIS CAPITULAR

Este capítulo presentó una breve introducción resumida del significado de la evaluación, sus antecedentes y distintos lazos disciplinares. Pretendía contextualizar e identificar sus vínculos disciplinares, en especial aquellos con las políticas públicas para generar alguna familiaridad. Entre otras cosas, se resaltó la autonomía disciplinar de la que goza y la importancia para las agencias gubernamentales y para los gobernados. Es de aclarar que la evaluación no nace con el estudio de las políticas públicas y que su funcionalidad es multidisciplinar, toda vez que se aplica a distintos campos como administración, economía, educación y, por supuesto, a la acción política y gubernamental. En ese sentido, se ilustraron los nexos con las ciencias de la administración, la educación, el desarrollo y la economía y, por supuesto, con las acciones gubernamentales, a fin de establecer las distintas aplicaciones de la evaluación en dichos dominios.

REFERENCIAS

- Anderson, J. E. (2006). *Public Policymaking. An introduction* (6a. ed.). Houghton Mifflin Co.
- Berner, M. y Bronson, M. (2005). A Case Study of Program Evaluation in Local Government: Building Consensus through Collaboration. *Public Performance & Management Review*, 28(3), 309-325. <http://www.jstor.org/stable/3381156>
- Carrillo, J. (2009). La dimensión política de la evaluación de políticas. *Opera*, 9, 53-65.
- CLACSO (2017). Teoría de la dependencia. <http://bibliotecavirtual.clacso.org.ar/ar/libros/clacso/crop/glosario/t.pdf>
- Coleman, J. S. (1988). Social Capital in the Creation of Human Capital. *The American Journal of Sociology*, 94, S95-S120.
- Davis, D. F. (1990). Do You Want a Performance Audit or a Program Evaluation? *Public Administration Review*, 50(1), 35-41.
- DeLeon, P. (1982). *Policy Evaluation and Program Termination*. The RAND Corporation Paper Series. The Rand Corporation.
- Ekelund, R. B., Jr. (1968). Jules Dupuit and the Early Theory of Marginal Cost Pricing. *Journal of Political Economy*, 76(3), 462-471.
- Friendly, M. (2007). André-Michael Guerry's Moral Statistics of France: Challenges for Multivariable Spatial Analysis. *Statistical Science*, 22(3), 368-399. <https://doi.org/10.1214/07-STS241>.
- Hogan, L. R. (2007). The historical development of program evaluation: exploring the past and present. *Online Journal of Workforce Education and Development*, 11(4), 1-14.
- Jones, C. O. (1969). The Policy Approach: An Essay on Teaching American Politics. *Midwest Journal of Political Science*, 13(2), 284-297.
- Joyce, L. (1980). Developments in Evaluation Research. *Journal of Occupational Behaviour*, 1(3), 181-190.

- Koontz, H. y O'Donnell, C. (1975). *Elementos de administración moderna*. McGraw-Hill.
- Lasswell, H. D. (1963). *El futuro de la ciencia política*. Trad. de Julio Cerón. Tecnos.
- Lasswell, H. D. (2000). La concepción emergente de las ciencias de las políticas, en: Luis Aguilar V. (ed.), *El estudio de las políticas públicas*. Miguel Ángel Porrúa.
- Lasswell, H. D. (2000). La orientación hacia las políticas, en: Luis Aguilar V. (ed.), *El estudio de las políticas públicas*. Miguel Ángel Porrúa.
- Morehouse, T. A. (1972). Program Evaluation: Social Research versus Public Policy, *Public Administration Review*, 32(6), 868-874. <http://www.jstor.org/stable/974647>
- Nakamura, R. (1987). The Textbook policy process and implementation research. *Policy Studies Review*, 7(1), 142-154.
- Ocampo, J. A. (2012). *La historia y los retos del desarrollo latinoamericano*. Comisión Económica para América Latina y el Caribe (CEPAL).
- Petty, W. (2004). *Essay on Mankind and Political Arithmetic*. The Project Gutenberg eBook # 5619. Transcribed from the Cassell & Co., editado por David Price.
- Rossi, P. H.; Freeman, H. E. y Lipsey, M. W. (1998). *Evaluation. A systematic Approach* (6a. ed.). Sage.
- Royse, D.; Thyer, B. y Padgett, D. (2016). *Program Evaluation: An Introduction to an Evidence-based Approach*. Cengage Learning.
- Taylor, F. W. (1987). *Principios de la administración científica*. Ateneo.
- Weiss, C. H. (1998). *Evaluation: methods for studying programs and policies*. Prentice Hall.

CAPÍTULO 2

Ámbitos, modalidades y propósitos

INTRODUCCIÓN

¿En dónde toma lugar la evaluación? ¿Cuántos tipos de evaluación existen? ¿Cuál es la finalidad de la evaluación? Este capítulo dará respuesta a estos interrogantes que conducen al contexto en el que toma lugar un ejercicio evaluativo, mismo que condiciona, constriñe y, en ocasiones, alienta su realización. En primera instancia, abordará los puntos dentro de un proceso donde se adelanta la evaluación, así como los campos donde ocurre, remitiendo así al mundo de las políticas públicas. Ofrece también una tipología regular, considerando las distintas posibilidades en función de su ocurrencia. Por último, identifica los distintos escenarios para efectuarla y los propósitos perseguidos.

2.1 LOS ESCENARIOS DE LA EVALUACIÓN

La evaluación está vinculada a cualquier fenómeno humano; sin embargo, en esta ocasión el foco de interés es, en general, el de las políticas públicas y, en particular, el de las políticas sociales o los programas gubernamentales. De forma implícita, los interrogantes giran en torno a cuál, cómo y para qué es la actividad gubernamental. Desde siempre, esta se ha organizado en programas de diversa clase y en distintos niveles, con los que se intenta incidir en el bienestar o cambiar comportamientos, actitudes, valores o conductas de los beneficiarios. También intenta cambiar las organizaciones e instituciones gubernamentales o las comunidades involucradas. La característica común de los programas del gobierno es su pretensión de mejorar las condiciones de vida de los receptores de su acción.

Los programas gubernamentales, entonces, son el centro de atención de la investigación evaluativa. Está implícita la idea de que la acción gubernamental importa y es definitiva para la sociedad, la economía y la política misma. Pero, ¿qué es un programa? Definido de manera simple, “es una colección organizada de actividades diseñadas para alcanzar ciertos objetivos” (Royse, Thyer y Padgett 2010, p. 5). Asimismo, la finalidad de un programa es la

resolución de un problema y, en consecuencia, se espera que tenga un impacto sobre los beneficiarios; los programas gubernamentales son emprendimientos complejos, pero a la vez una amalgama diversa (sueños y personalidades; oportunidades y teorías; insumos, estructuras orgánicas, clientes y actividades; presupuestos y grandes intenciones).

En la actualidad, los gobiernos se orientan por planes, programas, subprogramas y proyectos, instrumentos de planeación que soportan el desempeño del sector gubernamental en cualquiera de sus niveles. En estos instrumentos se asignan responsabilidades, se distribuyen presupuestos y se establecen mecanismos de control y rendición de cuentas. El marco institucional gubernamental opera para distintos sectores de la economía y de la política, permitiendo a sus interlocutores la identificación de límites y competencias, y facilitando las relaciones intergubernamentales y el control de los poderes complementarios.

La evaluación de la acción gubernamental por medio de sus programas tiene ya alguna tradición. En efecto, el gobierno federal de Estados Unidos, en un esfuerzo por identificar en los programas contra la pobreza, aquello que funcionaba y lo que no, en 1967 establece la evaluación como una función del sector público. Desde entonces, es requisito de un amplio rango de programas, pero en especial, los relacionados con asuntos socioeconómicos y originados en el paquete legislativo como parte de las iniciativas de la “gran sociedad” (Morehouse, 1972). Por lo que muchos gobiernos adoptaron esta modalidad de intervención para ordenar, focalizar y estructurar el gasto público de manera adecuada y propiciar el buen uso de los recursos públicos. La evaluación, de manera simple, intenta verificar qué tan buenos y efectivos son los programas establecidos, esto es fomentar y procurar buenos programas sociales.

Normativamente, un buen programa social es aquel que presenta ciertos elementos que le permiten lograr su alcance. Los componentes de un “buen programa gubernamental” incluyen la estructura del personal que fija o determina sus responsabilidades, la jerarquía y las capacidades para cumplir las funciones y aportar a los objetivos del programa (Royse, Thyer y Padgett *et al.*, 2010). A la vez, aunque pareciera evidente, el respaldo financiero es un elemento fundamental dispuesto bajo la modalidad de recursos presupuestales que permiten atender todas las necesidades y funcionar de forma adecuada. Adicional a la suficiencia presupuestal, la estabilidad de los recursos, aunque complementaria, está sujeta y se afecta por la dinámica y los vaivenes propios de la política. Por otra parte, los programas son visibles y reconocidos por el público; en consecuencia, tienen una identidad como también una filosofía de servicio, elementos que en últimas determinan cómo es y cómo debe ser tratado el grupo objetivo. Un buen programa se acompaña de un modelo teórico que le da consistencia y sirve como guía para la intervención pública. Más allá de las anteriores características, conviene agregar que, en esencia, un buen programa es aquel que cumple su finalidad en los hechos y que estos resultados están respaldados por la evidencia empírica y por los testimonios de aquellos que reciben sus beneficios. Conviene recordar que en la dimensión política hay tendencias a magnificar resultados como también a descalificar los logros. La subjetivi-

dad y los propósitos particulares pueden estar distorsionando la realidad de la acción gubernamental. En concreto y para los propósitos del presente trabajo, las políticas y los programas públicos son los escenarios naturales de la evaluación.

2.1.1 Condiciones favorables y aprendizaje

Es indudable que la consolidación de los gobiernos democráticos y abiertos y la emergencia de los programas gubernamentales han privilegiado la evaluación, haciéndola popular o conveniente y demandando la formación de especialistas. En consecuencia, los sistemas educativos han incorporado programas de posgrado para atender las demandas del mercado laboral por dichos especialistas. El ascenso de la evaluación se aceleró con los discursos de la Nueva Gestión Pública (NGP), al recomendar la introducción de técnicas gerenciales en los asuntos públicos y elevar así los niveles de desempeño y guiar los procesos de desarrollo.

La NGP ha propiciado nuevas posturas ciudadanas frente a la actividad pública y gubernamental que favorecen el control social y la evaluación del desempeño gubernamental. En efecto, el ciudadano es más propenso a exigir cuentas, escrutar la acción gubernamental y demandar responsabilidad frente a los recursos públicos. Es de destacar que la internacionalización y la globalización dinamizaron la cooperación internacional y las acciones de agencias multilaterales; por tanto, la ejecución de sus proyectos y agendas también demanda el apoyo de evaluadores expertos.

2.1.2 Las organizaciones en torno a la evaluación

El ascenso de la evaluación de las políticas se ratifica con el establecimiento de asociaciones, agencias y colectivos que agrupan a los expertos para la difusión de técnicas, formación de personal y promoción de la acción evaluativa. A continuación, una breve referencia de algunas instancias que contribuyen a la mejora continua, al adecuado desempeño de los gobiernos y a la consolidación de los principios democráticos. En primera instancia, se encuentran aquellas organizaciones que a nivel internacional destacan, bien por su tradición y su prestigio; después, se presentan iniciativas recientes de algunos países latinoamericanos que vienen consolidando esta función y animando la discusión en torno a los elementos teóricos de la investigación evaluativa.

1. Asociación Americana de Evaluación (*American Evaluation Association*). Una de las organizaciones más emblemática del ejercicio evaluativo. Desde su establecimiento (1986) tiene como misión mejorar las prácticas evaluativas y sus métodos mediante el fomento de su uso, su promoción como profesión y apoyo a la generación de la teoría y del conocimiento sobre la efectividad de la acción humana. La organización se fundamenta en valores como calidad, ética y sensibilidad cultural para el mejora-

miento de la función pública; asimismo, centra su actuación en los procesos de toma de decisiones, formulación de políticas y mejoramiento de programas. A la fecha cuenta con 25 afiliados y 7,000 profesionales de la evaluación, con los que comparte información y recursos, y promueve el desarrollo profesional.

2. Sociedad Británica de Evaluación (*UK Evaluation Society*). Organización de profesionales que vela por la promoción y el mejoramiento de la teoría, la práctica, el entendimiento y la difusión de la evaluación y su contribución al conocimiento público. Fundada en 1994, tiene afiliaciones de profesionales, practicantes, comisionados del nivel nacional y regional, comunidades académicas, consultores independientes y sector voluntario. Apoya la educación, el entrenamiento y, principalmente, representa a los interesados en la evaluación gubernamental y del sector público. Es miembro de la Academia de Ciencias Sociales del Reino Unido y mantiene vínculos con la Sociedad Europea de Evaluación y la Organización Internacional sobre la Evaluación.
3. Asociación Canadiense de Evaluación (*Canadian Evaluation Society*). Fundada en 1981, promueve el desarrollo profesional de sus afiliados a nivel nacional y regional, por medio de conferencias, competencias, publicaciones y cursos básicos, intermedios y avanzados. En 2009 fue la primera asociación en el mundo en crear un programa profesional acreditado presencial, pero solo hasta 2016 lo empieza a ofrecer a distancia. En la actualidad cuenta con cerca de 1,800 miembros.
4. Red de Monitoreo y Evaluación (M&E). Iniciativa impulsada en 2005 por el Banco Mundial, el Programa de Implementación del Pilar Externo del Plan de Acción a Mediano Plazo para la Efectividad del Desarrollo PRODEV del Banco Interamericano de Desarrollo (BID) y un creciente grupo de países de América Latina. Según la agencia mencionada, el PRODEV es una iniciativa para financiar cooperaciones técnicas y capacitación de funcionarios enfocados en fortalecer la efectividad de los gobiernos de la región de América Latina y el Caribe, a fin de que puedan alcanzar mejores resultados en sus intervenciones de desarrollo. El PRODEV cuenta con herramientas flexibles de apoyo, las cuales permitirán que algunos países logren el nivel deseado de gestión pública en menor tiempo que otros (CoPLAC GpRD, 2011).

En general, pretende mejorar el nivel de desempeño de las organizaciones públicas. La red, a su vez, intenta servir como escenario para el análisis de logros y dificultades durante la institucionalización de los sistemas de monitoreo y evaluación en los niveles nacional y subnacional. Asimismo, promueve el diálogo, comparte experiencias, fortalece la capacidad institucional y estimula la colaboración. Aspira a ser un espacio eficaz para fortalecer el conocimiento y las capacidades en monitoreo y evaluación, contribuyendo a las metas del buen gobierno y también a la reducción de la pobreza. En 2017 cambió a Comunidad de Profesionales y Expertos de América Latina y del Caribe en Gestión para Resultados en el Desarrollo (CoPLAC GpRD).

5. Consejo Nacional de Evaluación de la Política de Desarrollo Social (Coneval). Es importante reseñar los esfuerzos gubernamentales realizados en México para consolidar una cultura de la evaluación en el sector público. Desde 2004, el Coneval funge como “organismo público descentralizado de la administración pública federal con autonomía y capacidad técnica”. Genera también “información objetiva sobre la situación de la política social y la medición de la pobreza” en el país. Bajo estas consideraciones, regula y coordina “la evaluación de la política nacional de desarrollo social y todo tipo de políticas, programas y acciones ejecutadas por las dependencias públicas”. De igual modo, establece “los lineamientos y criterios para la definición, identificación y medición de la pobreza, garantizando la transparencia, objetividad y rigor técnico” (Coneval, *s/f*).

El Sistema Nacional de Desarrollo Social es “un mecanismo permanente de concurrencia, colaboración, coordinación y concertación del gobierno federal, de las entidades federativas y los municipales, así como los sectores social y privado” (Ley General de Desarrollo Social, 2004, p. 9). El artículo 72 de la mencionada ley determina que:

[...] la evaluación de la Política de Desarrollo Social estará a cargo del Consejo Nacional de Evaluación de la Política de Desarrollo Social, que podrá realizarla por sí mismo o a través de uno o varios organismos independientes del ejecutor del programa, y que tiene por objeto, revisar periódicamente, el cumplimiento del objetivo social de los programas, metas y acciones de la Política de Desarrollo Social [...] (p. 17).

A su vez, el artículo 81 establece que el Coneval,

[...] es un organismo público descentralizado, con personalidad jurídica, patrimonio propio, autonomía técnica y de gestión de conformidad con la Ley Federal de las Entidades Paraestatales. Tiene por objeto normar y coordinar la evaluación de las Políticas y Programas de Desarrollo Social, que ejecuten las dependencias públicas, y establecer los lineamientos y criterios para la definición, identificación y medición de la pobreza, garantizando la transparencia, objetividad y rigor técnico en dicha actividad (Ley General de Desarrollo Social, 2004, p. 18).

6. Sistema Nacional de Evaluación de la Gestión y Resultados Sinergia. Establecido en Colombia hace casi tres décadas pretende:

[...] mejorar la efectividad en la formulación y ejecución de las políticas del Plan Nacional de Desarrollo (PND); medir y evaluar oportunamente la gestión, la ejecución y el impacto de las principales políticas y programas gubernamentales; fortalecer el manejo gerencial de la inversión pública; generar la información adecuada para la asignación de los recursos, y mejorar las estrategias de rendición de cuentas a la ciudadanía para fortalecer el control social (Sinergia DNP, *s/f*).

Sinergia se ha consolidado en el país como fuente de información relevante en materia de seguimiento y evaluación de políticas públicas. Liderado por el “Departamento Nacional de Planeación, gestiona el manejo de la información relacionada con la agenda de evaluación, el seguimiento del plan nacional de desarrollo y el afianzamiento de la cultura de gestión pública por resultados en el país [...]” (Sinergia DNP, s/f). Fue la Constitución de 1991 la que introdujo la necesidad de implementar “un sistema de evaluación que midiera los resultados de las políticas y proyectos del gobierno”. Respondiendo así a la demanda de “modernizar el Estado, garantizar y mejorar la eficiencia de los recursos públicos, y democratizar la administración pública mediante un permanente y completo seguimiento y evaluación de las políticas públicas” (Sinergia DNP, s/f).

Es de anotar que la evaluación de la acción gubernamental se ha extendido por la región latinoamericana y se observan iniciativas en Argentina, Brasil, Bolivia, Chile, Ecuador, Perú, Paraguay y Uruguay. Esta evidencia corrobora el buen momento que atraviesa la evaluación de las políticas públicas.

2.2 MODOS DE EVALUACIÓN

Existen múltiples y variadas formas de evaluar, el tipo estará determinado por el momento en que ocurre, el enfoque, el número y la adscripción de los participantes y los propósitos, entre otros factores.

2.2.1 Según la finalidad

De acuerdo con la finalidad, tres son los tipos de evaluación: de insumos, de procesos y de resultados (Majone, 1997). Esta clasificación ha sido centrada en el momento en que se emprende dentro del ciclo de políticas o programas públicos. Para los insumos, se enfatizan los recursos, habilidades y personas disponibles. Se centra en la cantidad y la calidad técnica del personal, la información disponible, el nivel de financiamiento, el apoyo político, etc. En concreto, se estiman los resultados probables o la viabilidad de un nuevo proyecto. Para los procesos se consideran los métodos empleados para transformar los insumos políticos, económicos, etc., en productos o resultados. Especial énfasis merecen las operaciones y las actividades que permiten transformar los insumos para alcanzar resultados concretos (productos o servicios). En función de los resultados se valoran los productos o consecuencias de un programa particular se trata de un procedimiento obligado para estimar el valor de cualquier actividad deliberada. Asimismo, consiste en comparar las metas u objetivos con los resultados.

2.2.2 Según el momento

Por lo regular, los procesos evaluativos en función del momento se denominan *ex ante*, concomitante y *ex post*. La evaluación *ex ante* resulta útil y conveniente si se persiguen los propósitos de optimalidad (Todd y Wolpin, 2008). De esta manera, la aspiración es maximizar el impacto a un costo determinado, que es una alternativa a la evaluación al final del proceso de política pública. Asimismo, pretende corregir algunas debilidades de la evaluación final, ya que es imperativo conocer y evaluar con anterioridad, teniendo así la posibilidad de reaccionar frente a potenciales fracasos.

Esta evaluación admite el análisis de un programa previo a su puesta en marcha y los ajustes del caso para su mejora, sin comprometer su inversión (Todd y Wolpin, 2008). Nada elemental al demandar simulación de impactos potenciales y cálculo de tasas y coberturas con gran antelación, a fin de visualizar alternativas de mejoría. Para esto, es necesario adoptar una suerte de perspectiva experimental de tal manera que se puedan alterar algunos parámetros o variables y analizar sus consecuencias. Por obvias razones, este tipo intenta evitar gastos superfluos en programas que no prometen o anticipan resultados efectivos. La evaluación *ex ante* proporciona alguna evidencia de resultados e impactos para decidir con acierto coberturas, indicadores de evaluación y futuras estrategias de implementación. La modalidad *ex ante* puede asimilarse a la evaluación prospectiva, en la medida que intenta estimar posibles consecuencias de una propuesta o alternativas de políticas en competencia. Incluye también análisis lógicos, conceptuales y empíricos de lo que pudiera ocurrir (GAO, 1990). Permite tomar decisiones anticipadas sin comprometer recursos ni actividades.

La evaluación concomitante o intermedia examina los primeros resultados de las intervenciones y presenta las modificaciones necesarias para alcanzar los objetivos previstos. Lo anterior es apropiado cuando el programa está en curso y presenta alguna estabilidad, ya que permite analizar su pertinencia y coherencia con necesidades, objetivos, actividades y resultados propuestos a la luz de las circunstancias imperantes durante la ejecución. Asimismo, se valora la experiencia y el conocimiento acumulados durante la implementación de la política, por lo que deberá ofrecer una perspectiva de lo hecho y una anticipación de los resultados finales que se persiguen.

De igual modo, la evaluación intermedia sirve para ajustar estrategias, procedimientos y el plan en general cuando se anticipan dificultades o desajustes en los atributos de la política; indaga también sobre las condiciones que pueden facilitar o impedir el logro de los objetivos. Con esta valoración se verifica lo que sucede con la planeación, el diseño, la ejecución y los resultados del programa o política para intervenir de forma oportuna. Asimismo, se aprecia un punto de coincidencia con el monitoreo y con la función de control, al punto que se confunden por adelantarse durante la ejecución y por tener la finalidad de emprender medidas correctivas para mantener el espíritu de la iniciativa.

La evaluación *ex post* o *a posteriori* es animada por los mandatos legales y la teoría económica para determinar los beneficios, resultados e impactos de las políticas o iniciativas de ac-

ción pública. Es de reconocer, la compatibilidad entre los propósitos de esta valoración y las previsiones legales como las consideraciones utilitaristas (Harel, Safran y Segal, 2005). Se trata de establecer si la acción gubernamental, mediante sus políticas, ha otorgado los beneficios a individuos cuyo bienestar estuviera en riesgo y, de paso, sopesar si dichos beneficios corresponden a las inversiones realizadas, si cumplieron lo prometido, si hay cabida a sanciones en caso de incumplimiento o si brindaron la máxima cobertura.

Dado que estimaciones *ex ante* resultan extremadamente limitadas para la evaluación científica, la valoración *ex post* se reconoce como más apropiada para la evaluación de las políticas (Herrick y Sarewitz, 2000). Es considerada como constructiva y realista al emprenderse sobre información histórica y no sobre estimaciones y predicciones que podrían resultar problemáticas en ambientes políticos y fiscales inescrutables. Bajo esta perspectiva, la evaluación *ex post* apoyaría de manera importante los dilemas de la investigación evaluativa, el aprendizaje social y la perspectiva adaptativa de la gerencia pública.

Esta evaluación se realiza una vez finalizada la política mediante la recolección de información acerca de las actividades, resultados o impactos obtenidos por los beneficiarios de una política o programa particular durante un periodo determinado. Constituye también un ejercicio para establecer no solo el éxito final de la intervención a la luz de los propósitos iniciales, sino también el aprendizaje obtenido durante dicho periodo.

2.2.3 De acuerdo con sus enfoques

Como cualquier otra actividad científica, la evaluación también expresa sus preferencias hacia las diversas tradiciones ontológicas y epistemológicas. Lo anterior, por supuesto, tiene implicaciones metodológicas. Así, puede apreciarse que los juicios pueden tener naturaleza cuantitativa o cualitativa, según la naturaleza de la información. La evaluación cuantitativa, vinculada al positivismo, trata con números, mientras que la cualitativa se relaciona con palabras. Lo cuantitativo tiene una visión estructurada y está relacionado con la perspectiva hipotética-deductiva, esto es parte de la teoría y quiere verificar la validez de esta, en su caso de estudio, mediante la validación de la hipótesis. Asimismo, tiene como punto de partida la variable y esta es parte constitutiva de su forma de operar y su propósito, que en sus inicios consistía en medir.

Por tanto, la variable es una idea central en la evaluación cuantitativa. De manera simple, el término invoca algo que varía, bien en cantidad, intensidad o monto, según la unidad de análisis que se considere. Las variables, por tanto, “son propiedades, características o atributos que se dan en grados o modalidades diferentes en las personas y, por derivación, en los grupos o categorías sociales” (Carrillo, 2020, p. 12). Son variables: la edad, el ingreso, la educación, el género, la ocupación, etc. Pueden clasificarse según diversos criterios, a saber: según la relación, la escala o su naturaleza. En consecuencia, la evaluación cuantitativa compila datos que pueden ser transformados numéricamente, de tal manera que el análisis está

basado en las técnicas estadísticas. Los datos se recogen mediante instrumentos estandarizados para sistematizarlos y administrarlos, así como para generar y analizar información. Los datos de encuestas, cuestionarios, registros o puntuaciones se transforman en series de valores numéricos. Para su análisis, se recurre a métodos estadísticos y para su reporte a tablas, gráficas y modelos acompañados con un texto verbal que explica el significado de los resultados (Neuman, 1997).

Las técnicas para el análisis cuantitativo pueden llegar a ser extensas y cada día más sofisticadas. Asimismo, pueden ser utilizadas para describir el comportamiento de las variables de manera aislada (análisis univariado) o para estudiar las relaciones que existen entre las variables (análisis bivariado). La naturaleza de las hipótesis tiende a estudiar la causalidad y por ello proponen la relación entre dos o más variables. Una hipótesis es una proposición tentativa acerca de los fenómenos que se estudian y son validadas por medio de distintas técnicas de la estadística. Para la evaluación cuantitativa, la confiabilidad y la validez son dos condiciones centrales. Confiabilidad indica que la repetición de varias pruebas, procedimientos o experimentos llevan siempre a los mismos resultados. La validez, por su parte, habla de los indicadores y su correspondencia con lo que intentan medir. En cambio, el análisis multivariante se caracteriza por asumir una posición más integral o compleja de los fenómenos y sus técnicas se encuentran en expansión, gracias al desarrollo e impacto de la informática. El análisis multivariante refiere todos los métodos estadísticos que analizan simultáneamente las medidas múltiples de cada individuo u objeto sometido a estudio (Hair, Anderson, Tatham y Black, 1999). En sentido estricto, son extensiones del análisis univariado y bivariado.

La perspectiva cualitativa es objeto de permanentes críticas por aquellos que dudan de su validez para el estudio y la evaluación de los fenómenos sociales. Por lo que puede afirmarse que no gozan de popularidad, dada la complejidad de sus técnicas y de las dificultades para comunicar sus hallazgos; quizá esos sean sus principales desafíos. Aun así, dada la expansión de la computación, sus técnicas están siendo acogidas por algunas comunidades académicas, quienes empiezan a valorar la potencia para la descripción y el análisis inferencial de la información disponible en la actualidad.

Por su parte, lo cualitativo deriva de una posición filosófica distinta de lo cuantitativo. Como se ha reiterado en otros capítulos, lo primero que se niega desde aquí es que exista una realidad externa y única, como lo afirma el positivismo. La interpretación es la finalidad de la investigación y seguramente de la evaluación, y echa mano de las técnicas que ello permiten. La información tiene naturaleza distinta y difiere en mucho del dato y la cifra. La información proviene del trabajo de campo que realiza el evaluador mediante la investigación etnográfica, al estilo del realizado por los antropólogos que pasan largos periodos en el campo para conocer y comprender la cultura, los modelos mentales y las actividades de los nativos; el énfasis está en comprender, más que en explicar. Las comparaciones entre grupos no están en su agenda y, por tanto, están soportadas en el detallado conocimiento del proceso y de sus efectos sobre los participantes y no en comparaciones estadísticas. El evaluador analiza los

datos tal como emergen, para después desarrollar fases de medición como respuesta al conocimiento adquirido.

Los estudios cualitativos comparten, por lo menos, tres ingredientes clave, la medición naturalística (en el sitio en que ocurren los hechos), el involucramiento del investigador y su cercanía con el sujeto de estudio y los reportes que revelan densas narrativas (Gilbert, 2001). Es importante aclarar que la evaluación cualitativa no se asimila a ningún estudio de opinión (la popular percepción), ya que sus métodos distan en mucho, pues estos son empíricos y sistemáticos, además de que se basan en la documentación y el análisis de los datos.

¿Cuándo conviene lo cualitativo y cuándo lo cuantitativo?, esta es otra consideración que debe resolverse. Las estrategias cuantitativas son adecuadas cuando se disponen datos de una muestra grande y cuando el tiempo es limitado. Sin embargo, esta no ofrece descripciones vívidas de los sujetos estudiados. Mientras los números y el análisis estadístico valoran el grado de funcionamiento de un programa, los estudios individuales cualitativos permiten entender la forma como se logró dicho funcionamiento; también generan una mayor conciencia de la perspectiva de los participantes del programa y, a menudo, una mayor sensibilidad de sus intereses. En general, lo cualitativo permite entender la dinámica del desarrollo y evolución del programa, así como una mayor conciencia sobre el tiempo y la historia de la acción emprendida. A la par, otorga una mayor conciencia del contexto y la habilidad para entrar en la escena del programa, para ser parte de este sin preconcepciones y, en general, lo cualitativo otorga mayor flexibilidad a las perspectivas.

A la luz de la distinción de la primera tipología de evaluación (insumos, procesos y resultados), los métodos cualitativos están asociados a la evaluación de procesos (estudiar el cómo y el porqué del programa). También son útiles a la evaluación formativa para mejorar el programa antes de su implementación y son prioritarios cuando se trata de estudiar aspectos difíciles de capturar y medir (la moral del *staff*; la toma de decisiones de los ejecutivos, los malentendidos a través de las culturas, la percepción de los clientes, etc.). La ventaja de los enfoques cualitativos radica en que permiten examinar fenómenos complejos, sin depender de una recolección estructurada demandada por los estudios cuantitativos. Otra ventaja de la perspectiva cualitativa es el contacto directo y personal del evaluador y los actores de la evaluación. De paso, conviene resaltar que la visión cualitativa tiende a empoderar a los participantes.

2.2.4 De acuerdo con la intensidad y profundidad

La dicotomía formativa-sumativa es una posibilidad que, por lo regular, se presenta a la hora de hacer las distinciones en una tipología de las evaluaciones. Aunque su origen se encuentra en la práctica educativa, dicha posibilidad puede vincularse al ejercicio evaluativo de las políticas. Lo primero que habría que señalar es que lo formativo tiene un interés en la evaluación de procesos y de resultados. La formativa opera cuando ha transcurrido poco tiempo

desde el diseño e implementación de la política, por lo que es muy probable que sus operadores reconozcan la existencia de algunas disfuncionalidades en el funcionamiento, dada su reciente ejecución. No desconocen la presencia de quejas que demandan una revisión; por tanto, el evaluador es convocado para obtener una retroalimentación constructiva del programa. Bajo este tipo de valoración, el interés es la mejora del programa y no tanto su desaparición.

Se recurre a las evaluaciones formativas cuando se desea ajustar y mejorar las intervenciones públicas. Sirven más como guías y directrices de los programas, por lo que una buena analogía es la de un instructor experto en conducción, sentado al lado de un conductor principiante. Le enseña a ser más habilidoso, observando el proceso y haciendo sugerencias constructivas; por tanto, el instructor está más interesado en el proceso que en el destino final. El evaluador formativo podría mirar las interacciones entre beneficiarios y operadores, las estrategias de la dirección, la filosofía de servicio y los costos del programa, entre otros criterios (Chelinsky, 1985). Esto, de alguna manera, se asemeja a la tan conocida técnica del *coaching*, pues sirve para determinar si un nuevo programa o un programa piloto ha sido implementado en los términos que fue planeado.

Esta evaluación identifica los servicios suministrados, su proveedor, tiempos, frecuencia y las circunstancias de la política, asuntos estos que llevan a contextualizar la evaluación al interior de las agencias operadoras. Por tanto, se reconocen sus fortalezas y debilidades, se recopila información entrevistando al staff o a los beneficiarios o revisando documentos de avances e informes. La evaluación formativa está interesada y comprometida con los procesos durante la implementación del programa o política (Beer y Bloomer, 1986). Es útil para adaptar determinadas intervenciones a grupos específicos y revisar estrategias o diseños.

La decisión por la evaluación formativa depende del programa, de las preferencias del evaluador, de la agencia operadora y del contexto en que se demande. Tres son las formas para emprenderla:

1. Localizar modelos estándares que sirvan como referentes (*benchmarking*) de comparación, a fin de reconocer discrepancias o identificar otros niveles de gobierno o agencias que realizan actividades similares.
2. Buscar ayuda de expertos de reconocida reputación y con cierta autoridad para que soporten el proceso.
3. Conformar un comité para orientar la actividad. En ocasiones, la evaluación formativa se asimila al monitoreo de un programa, dado su carácter rutinario, inicia revisando las metas y objetivos específicos del programa y los compara con otras agencias de servicio. Se trata también de determinar si el programa sirve a los fines y a los beneficiarios establecidos en la etapa de diseño.

Por su parte, la evaluación sumativa está interesada en los logros agregados que una política ha podido acumular, de ahí su denominación, la cual se efectúa, por tanto, una vez que el programa haya terminado, concentrándose en esencia en los resultados y en los impactos. Esta valoración se basa en un conocimiento empírico de los resultados finales de un programa y busca respuestas en torno a la satisfacción de sus beneficiarios, el logro de metas y las mejoras alcanzadas. Medir los resultados y su impacto puede considerarse el propósito central de este tipo de evaluación. Este ejercicio proporciona a los tomadores de decisiones, así como información sobre la efectividad del programa y puede conducir a decidir sobre su continuidad o desaparición, conveniencia de expandirlo a otras localidades o, por el contrario, recortarlo.

Desde el punto de vista metodológico, la sumativa se basa, en buena parte, en los enfoques cuantitativos para medir los resultados finales. Asimismo, puede tener interés en la efectividad, la eficiencia o el impacto, ya sea de manera individual o conjunta. Quienes practican la sumativa están interesados en el uso de técnicas que permitan la medición, el análisis estadístico, la prueba de hipótesis e incluso cualquier tipo de experimentos. Estas técnicas otorgan formalidad y sofisticación, pero exigen recursos. En efecto, tales modalidades requieren información sistemática, soporte técnico especializado y uso de tecnología; todos estos requerimientos imponen elevados costos de evaluación, los cuales difícilmente son asumidos. Las evaluaciones de impacto están dentro de las sumativas y exigen mayores recursos. Aun así, esto no debe ser una limitación, pues toda intervención requiere determinar los principales propósitos en cuanto a efectividad, eficiencia e impacto.

Las modalidades, la formativa y la sumativa de la evaluación llevan a pensar sobre sus focos e intereses que se persiguen. Al decir de algunos, lo formativo se preocupa por la implementación, el aprendizaje que se va adquiriendo y los ajustes que se pueden y deben realizar, de ahí su asociación con el monitoreo. Lo sumativo estudia los procesos acumulados y los agregados de una intervención, por lo cual se vincula al juicio global de una intervención. Aunque por lo general se adelantan de manera independiente, es necesario integrar la valoración de todas las intervenciones de desarrollo tanto en la formativa como en la sumativa (NU, 2010).

2.2.5 Según los actores

La evaluación participativa es una de las posibilidades explorada con mayor intensidad, la cual considera la inclusión de juicios de quienes participan en el programa: administradores/operadores, interesados, responsables o beneficiarios directos. En otras palabras, involucra a todos los *stakeholders* en las tareas evaluativas, bajo la premisa de que se amplía el espectro y el juicio de las acciones emprendidas y de los procesos y resultados de programas y políticas. Estas ideas tienen un fuerte arraigo con las nociones de la democracia, la perspectiva de los involucrados y el control social, quedando pendiente resolver las formas, los procesos, las calidades y los alcances de participación.

La evaluación participativa aspira, sobre todo, a animar a que los actores, desde sus distintos roles, superen la condición pasiva y asuman funciones de planeación, recolección, análisis y diseminación de los ejercicios evaluativos. Esto promueve el capital social, la transparencia y la responsabilidad de los gestores del programa y de los encargados de su evaluación. Por lo anterior, son muchas las ventajas de la evaluación participativa, sobre todo en momentos en que el ciudadano reclama mayores oportunidades para ejercer sus derechos y para la construcción del espacio público. De manera similar, dicha participación promueve la superación personal y fomenta la educación colectiva, facilitando la transferencia del conocimiento de los expertos hacia las comunidades locales (Schuftan, 1996 en Dobbs y Craig, 2002).

Aunque con algunas reservas y precauciones, prevalece la esperanza de que las comunidades se apropien de sus problemáticas y necesidades, se sumen a la propuesta de soluciones y sean sujetos de la política. Aun así, persisten condicionantes y restricciones para que la evaluación participativa sea de las más populares y exitosas. Es necesario adelantar estrategias para que su adopción sea menos traumática e institucionalizada sin mayor reticencia.

2.2.6 Según la adscripción del evaluador

La ubicación del evaluador da origen a la evaluación interna y externa afectando los propósitos del estudio y con hondas repercusiones para la confianza administrativa, la objetividad, la comprensión del programa, la utilización futura de los resultados y la autonomía. Esta distinción está mediada por diferentes circunstancias como la premura, las competencias institucionales, la disposición de recursos y el soporte humano, entre otras. Ambas posibilidades ofrecen algunas ventajas y limitaciones que bien vale la pena considerar. Quienes favorecen la interna privilegian el conocimiento de las reglas de operación, los antecedentes y la familiaridad con el objeto de la evaluación. Quienes aprueban la objetividad y la imparcialidad como atributos necesarios conceden mérito a la evaluación externa. Esta es iniciada y financiada por otras fuentes ajenas a la administración y, sin duda, emprendida para supervisar la implementación (Wollmann, 2007). La autoevaluación o evaluación reflexiva es una posibilidad que se viene promoviendo, por sus ventajosas posibilidades, en el ámbito educativo y gubernamental. La autoevaluación, en efecto, permite que tanto los actores como las organizaciones inicien procesos de reflexión acerca de su actuar a lo largo de un proyecto, programa o política.

En realidad, este tipo de evaluación constituye una nueva estrategia impulsada desde distintos espacios, con la idea de superar algunas limitaciones de las tradicionales modalidades y para afianzar la cultura de la evaluación. Así, por ejemplo, la autoevaluación participativa promovida por Naciones Unidas, por medio de algunas de sus agencias, propone que se lleven a cabo mediante ejercicios estructurados con la participación de la cooperación, los gobiernos, las agencias y, de ser el caso, los beneficiarios para evaluar en conjunto la efectividad,

la eficiencia, el impacto y la sostenibilidad de sus proyectos. Asimismo, puede tratarse de evaluaciones intermedias o finales, dependiendo de las necesidades, que buscan conocer sobre las lecciones aprendidas y consolidar las futuras implementaciones. “La autoevaluación debe ser una herramienta de aprendizaje y gerencial, al decir del secretario de las Naciones Unidas. Debe, en complemento, asegurar la calidad y ser un medio para expandir la responsabilidad y la rendición de cuentas de los estados miembros” (UN, 2018). La inclusión, como elemento central de esta modalidad evaluativa, atina a quitar el monopolio tradicional de las élites administrativas, incorporando actores de todos los niveles que puedan aportar al mejoramiento de las iniciativas, pues son ellos quienes conocen a detalle la esencia de las intervenciones y las perspectivas de los beneficiarios.

2.3 ESPÍRITU DE LA EVALUACIÓN

Vale la pena ratificar que la evaluación es parte de un proceso gerencial (táctico o estratégico), es decir, ayuda a quienes toman decisiones. Contribuye al mejoramiento de un programa, verifica su eficacia, el logro de propósitos o metas, la pertinencia de los procedimientos para su implementación y el uso adecuado de los recursos. En otras palabras, verifica la operación y los resultados del programa. La evaluación es un proceso de confrontación de estándares o patrones que estima el mérito del programa, comparando la evidencia del proceso y sus resultados con el conjunto de expectativas. El evaluador espera que las autoridades usen los resultados derivados para tomar decisiones acertadas.

Hay que reconocerlo, la evaluación, al igual que la adopción de un programa, genera expectativas de distintos actores, beneficiarios, donantes, responsables e incluso los mismos profesionales y académicos, que cultivan la disciplina de las políticas públicas. Por tanto, la función evaluadora tiene que ver con verificar el cumplimiento de dichas expectativas a partir de preguntas relacionadas con el programa o la política: ¿alcanzó sus metas?, ¿hubo algún grado de [dis]satisfacción?, ¿surgieron asuntos adicionales para la agenda pública? La función heurística es la función más reconocida de la práctica evaluativa; al evaluar se genera la información y el conocimiento necesarios para aprender, como también para innovar y contribuir al logro de los fines del sistema. La heurística se define como la capacidad de un sistema para el descubrimiento y la invención a la hora de resolver problemas usando la creatividad y el pensamiento lateral. El gran supuesto es que la experiencia puede contribuir a la solución de problemas, en la medida que sean similares. Y es que con su ejercicio se gana conocimiento, aprendizaje y un resultado deseado. El ensayo y error podría describir, de alguna manera, esta función, pues la experiencia acumulada en un proceso evaluativo privilegia la generación de conocimiento valioso. En general, el proceso integral de las políticas públicas (diseño, implementación y evaluación) se reconoce por su función heurística (Kulaç y Özgür, 2016).

Más allá del proceso de aprendizaje, existen otras motivaciones para emprender la evaluación de un programa. Subyacen motivos racionales, pero también políticos; por esto, en ocasiones se considera como subterfugio (Weiss, 1998). Sin lugar a duda, hay motivos diversos, uno de ellos es usarlo como cortina para ocultar verdaderas intenciones, en lugar de enfrentar una situación problemática. Lo usual es conformar un comité que, a su vez, comisiona una evaluación, ganando así tiempo y dilatando una decisión. De manera similar, cuando el gestor no quiere asumir la responsabilidad por las decisiones, en especial cuando hay facciones enfrentadas y no quiere comprometerse con alguna de ellas, ya que asume el ejercicio evaluativo como una fuente de apoyo. En otras ocasiones, se asume la evaluación como una fuente de legitimidad para decisiones tomadas con antelación; en este caso, se ocultan responsabilidades en un ejercicio evaluador. Cuando los resultados son positivos, es posible percibirla como una fuente de autoglorificación, en la que el administrador no desaprovecha la ocasión para hacerse visible y figurar. Por el contrario, cuando los resultados no son favorables, la emplea para justificar las debilidades, seleccionando aquellos aspectos superficiales que le favorecen.

En otro escenario, es posible usar la evaluación para formalizar algunos requisitos. A menudo, la exigencia es externa al programa y apenas comprensible, ya que los proveedores de los recursos quieren tener certezas. Las relaciones intergubernamentales, por ejemplo, establecen como mandato la evaluación de programas y proyectos financiados. Asimismo, se adopta la evaluación como medio para obtener apoyo financiero y para dar cumplimiento a una formalidad.

El ejercicio evaluativo también se reconoce como medio para implementar medidas correctivas. Directores y administradores quieren saber lo que pasa con el programa para corregir asuntos en caso de ser necesario; también les permite determinar, con la debida anticipación, cambios. La evaluación debe reconocerse como un momento en el que se evidencian asuntos experimentados que generan sabiduría y que se fortalece, comunalmente, al revelar debilidades y fortalezas, y que, en todo caso, marca el camino para seguir creciendo. Asimismo, se trata de una etapa que enseña, forma y estimula (aprendizaje organizacional). Esto revela que el espíritu de la evaluación no se orienta, en exclusiva, hacia lo instrumental de la acción, pues también se centra en la recopilación y el registro de la historia del programa, bajo el supuesto de que la historia enseña.

La evaluación, como medio para la rendición de cuentas a los *stakeholders*, es reciente; en especial, cuando condiciona o acota el futuro de los apoyos. Dadas las actuales condiciones presupuestales y las nuevas actitudes ciudadanas, las agencias públicas deben presentar reportes detallados de lo alcanzado con los fondos recibidos. Por tanto, la presión por la rendición de cuentas lleva a una evaluación sistemática. Por último, como medio para entender la intervención social, proporciona una oportunidad para estudiar sus consecuencias, teorizar sobre las iniciativas del programa, estudiar estímulos que conduzcan al cambio y condiciones ambientales en las que puede tomar lugar.

2.4 SÍNTESIS CAPITULAR

Una vez revisados los escenarios, las modalidades y los propósitos de la evaluación de los programas gubernamentales pueden apreciarse también, de manera precisa, las dimensiones y alcances de dicha función. Este apartado ha develado las diversas posibilidades disponibles para llevar a cabo un ejercicio evaluativo, con distinta naturaleza de la información, sus participantes y sus propósitos. Ello sugiere que el evaluador tiene que identificar la modalidad adecuada y estar en posibilidad de aplicar las técnicas que tales decisiones conllevan. Por otra parte, se identificaron razones que subyacen en cualquier evaluación, alertando acerca de las posibles intenciones que podrían ocultar o de las razones legítimas que la soportan. En ese sentido, pudiera firmarse que la evaluación de los programas traspassa lo técnico para incluir lo estratégico, lo social y lo político.

REFERENCIAS

- Beer, V. y Bloomer, A. C. (1986). Levels of Evaluation. *Educational Evaluation and Policy Analysis*, 8(4), 335-345.
- Carrillo, J. (2020, 25 de julio). Sesión sincrónica 1 [presentación de diapositivas]. Campus en línea, Unitec.
- Chelimsky, E. (1985). Comparing and Contrasting Auditing and Evaluation. Some notes on their relationship. *Evaluation Review*, 9(4), 483-503.
- Chelimsky, E. (1987). "What Have We Learned about the Politics of Program evaluation?" *Educational Evaluation and Policy Analysis*, 9(3), 199-213.
- Clarcke, A. (2001). Research and Policy-making Process, en: *Researching Social Life*, Gilbert, Nigel (ed.). Sage.
- Consejo Nacional de Evaluación de la Política y el Desarrollo Social (2019). <https://www.coneval.org.mx/quienessomos/Conocenos/Paginas/Quienes-Somos.aspx>
- CoPLAC-GpRD. Comunidad de Profesionales y Expertos en Latinoamérica y el Caribe en Gestión para Resultados en el Desarrollo. https://www.coplac-gprd.org/index.php?option=com_content&view=category&id=5&layout=blog&Itemid=55&lang=es
- Dobbs, L. y Moore, C. (2002). Engaging Communities in Area-based Regeneration: The Role of Participatory Evaluation. *Policy Studies*, 23(3/4), 157-171.
- Harel, A.; Safran, Z. y Segal, U. (2005). Ex-Post Egalitarianism and Legal Justice. *Journal of Law, Economics, & Organization*, 21(1), 57-75.
- Herrick, C. y Sarewitz, D. (2000). Ex-Post Evaluation: A More Effective Role for Scientific Assessments in Environmental Policy. *Science, Technology, & Human Values*, 25(3), 309-331.
- Kulaç, O. y Özgür, H. (2017). An Overview of the Stages (Heuristics) Model as a Public Policy Analysis Framework. *European Scientific Journal*, edición especial, 144-157.
- Majone, G. (1997). *Evidencia, argumentación y persuasión en las políticas*. Fondo de Cultura Económica.

- Morehouse, T. A. (1972). Program Evaluation: Social Research versus Public Policy, *Public Administration Review*, 32(6), 868-874.
- Naciones Unidas. Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura, Unesco (2010). La lente de la educación para el desarrollo sostenible: una herramienta para examinar las políticas y la práctica. La educación para el desarrollo sostenible en acción. Instrumentos de aprendizaje y formación, 2.
- Neuman, L. (1997). *Social Research Methods. Qualitative and Quantitative Approaches*. Allyn & Bacon.
- Royse, D.; Thyer, B.A.; Padgett, D. y Logan, TK. (2010). *Program Evaluation: An introduction*. Cengage Learning.
- Sistema Nacional de Evaluación de Gestión y Resultados, SINERGIA (s/f). Departamento Nacional de Planeación República de Colombia. <https://sinergia.dnp.gov.co/Paginas/Internas/Sinergia.aspx>. Consultado el 3 de marzo de 2019.
- Todd, P. E. y Wolpin, K.I. (2008). Ex-Ante Evaluation of Social Programs. *Annales d'Économie et de Statistique*, Ns. 91/92, Econometric Evaluation of Public Policies: Methods and Applications, 263-291.
- United Nations. United Nations Conference on Trade and Development, UNTACD (2018). Participatory Self-Evaluation. A Guide for Project Managers. Ginebra: The Evaluation and Monitoring Unit (EMU).
- United States of America, General Accounting Office (1990). *Prospective Evaluation Methods. The Prospective Evaluation Synthesis*. GAO/PEMD-10.I.I.O.
- Weiss, C. (1998). *Evaluation*. (2a. edición). Prentice Hall.
- Wollmann, H. (2007). Policy Evaluation and Evaluation Research, en: *Handbook of Public Policy Analysis. Theory, Politics, and Methods*, Fischer Frank, Gerald J. Miller y Mara S. Sidney (eds.), 393-402.

CAPÍTULO 3

Proceso de evaluación

INTRODUCCIÓN

¿Cómo dar comienzo a una evaluación? Este es tal vez uno de los interrogantes esenciales para quienes se inician en esta profesión. El proceso de evaluación es un asunto complejo que requiere planeación. Subyace la idea de que dicha función tiene varios momentos y un conjunto de elementos que, de principio a fin, permitan juzgar una acción o un programa. Los momentos evaluativos, a saber, son planeación (diseño), implementación y difusión de los hallazgos.

Un diseño riguroso de una evaluación, en lo esencial, reduce la probabilidad de desviación en los objetivos perseguidos. De igual manera, permite identificar los actores involucrados, la preparación para la recolección de datos e información apropiada para estudiar e identificar rutas alternativas y la estimación de tiempo y costos, y gestionar eventualidades. El plan reduce la incertidumbre ayudando a dar un sentido claro de orientación de la evaluación, pero, en todo caso, exige fortalezas en investigación social.

3.1 PROCEDIMIENTO GENERAL DE LA EVALUACIÓN

Por lo regular, los ejercicios evaluativos de las actividades gubernamentales han estado acompañados de manuales, guías y documentos especializados, mismos que intentan asistir a los involucrados (Wall, 2004; Meyer, 2015). A la par, las agencias multilaterales y las organizaciones no gubernamentales han animado esta actividad (Tearfund, 2001; UNDP, 2009). En los materiales mencionados suelen identificarse las etapas bajo una secuencia de actividades que deben adelantarse para abordar el proceso de principio a fin: Wall (2004) con nueve pasos, *General Accounting Office GAO* (1991) con cinco, *National Science Foundation* (2002) con seis y el Comité Conjunto de Sistemas de información (*Joint Information Systems Committee*) (2007) con seis pasos. Proponer preguntas de evaluación; establecer un diseño de la evaluación; acopiar información necesaria y analizarla; reportar y diseminar hallazgos e

implementarlos, son los elementos comunes de las distintas propuestas. Complementan el proceso de evaluación: desarrollar modelos conceptuales sobre el programa en cuestión; identificar *stakeholders* y determinar estrategias de trabajo.

Conviene insistir que la planeación de una evaluación merece especial atención por incidir en los resultados, los costos, los tiempos y la incertidumbre que encarna la actividad. La planeación, además de determinar la información requerida, los métodos factibles, las técnicas adecuadas y demás detalles de un proceso complejo, reporta el estado actual de un programa y las posibilidades de mejora del desempeño. Los esfuerzos se centran en el diseño de un plan para conducir una evaluación (Meyer, 2015) y sus características dependerán del tipo de evaluación, la clase de información requerida o disponible y los métodos a utilizar. Asimismo, las acciones se resumen en tres elementos: planeación, implementación y utilización (figura 3.1). El primer elemento corresponde al momento en que se determinan las concepciones y definiciones propias de la política pública y los criterios que guiarán la evaluación. La implementación lleva a la ejecución de todas las decisiones y se complementa con la recolección de información, identificación e incorporación de los *stakeholders*. La utilización, por último, determina cómo serán usados los hallazgos para el mejoramiento de las acciones en el marco de la política de referencia, a fin de mejorar su desempeño.

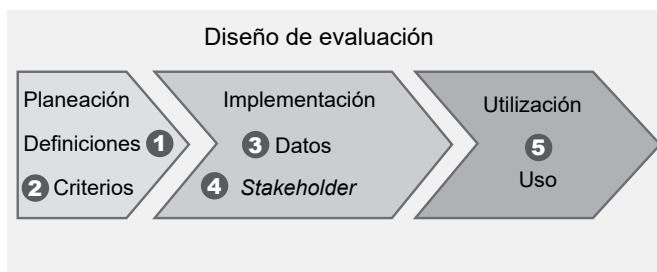


FIGURA 3.1. DISEÑO DE EVALUACIÓN

Fuente: elaboración propia con base en Meyer, 2015, p. 4.

3.2 EL PROCEDIMIENTO DE LA EVALUACIÓN

La evaluación no es una actividad compacta e inflexible. Quienes han estudiado dicha actividad encuentran que existen distintas etapas de evaluación que, de principio a fin, permiten llegar a un juicio objetivo del actuar de un programa o política. En algunas ramas de las ciencias sociales, la secuencia de actividades para evaluar es asimilable a un procedimiento investigativo. Tomando como referencia el documento del Comité Conjunto de Sistemas de Información (*Joint Information Systems Committee*) del Reino Unido, procede lo siguiente:

1. Identificación de *stakeholders*. A fin de entender el poder, posición y perspectivas de los distintos individuos y las organizaciones que tienen algún interés en una política particular o que pueden ser afectados por una eventual reforma (Buse *et al.*, 2005, citado en Nash, Hudson y Luttrell, 2006, p. 26). En algunas circunstancias, los *stakeholders* pueden cobijar comunidades y público en general. Los *stakeholders* tienen así, distinta naturaleza e intereses y recursos diversos, y el evaluador con su equipo tienen una interacción permanente durante dicho proceso. Unos pueden ser beneficiarios de las acciones del programa; otros, responsables en las distintas funciones administrativas e incluso pueden comprender donantes y patrocinadores del mismo programa. Su identificación y análisis permiten entender particulares reacciones o incidencias, procediendo una clasificación de sus intereses en el asunto y su cuantificación. A la vez, es necesario medir sus recursos de poder e impacto para fomentar o bloquear cualquier iniciativa o política.
2. Descripción y entendimiento de la política. Aunque parezca una verdad de Perogrullo, para evaluar una política es necesario entender a cabalidad lo que intenta lograr, caracterizando el contexto y las condiciones que la originaron para complementar el análisis de los grupos de interés. Es importante, a la vez, conocer las reglas que determinan el funcionamiento de la política, los procedimientos y los productos generados. Juzgar sin conocer al detalle todos los antecedentes, insumos, procedimientos, contexto cultural y demás particularidades, resultaría un ejercicio impreciso, sin legitimidad y con pocas probabilidades de éxito.
3. Diseño de la evaluación. El éxito de la evaluación lo determina, en buena parte, su adecuado diseño. Para ello es necesario producir un esquema o estrategia que ofrezca respuestas a interrogantes de carácter descriptivo, normativo y de impacto (causa-efecto), mediante elementos especificados con claridad. Las respuestas descriptivas deben proporcionar información acerca de las condiciones o eventos específicos del programa o la política (GAO, 1991). Este diseño debe precisar, por ejemplo, quién o quiénes están recibiendo el beneficio o, si es el caso, determinar los costos de un equipo o instrumento, etc. Por otra parte, la información de tipo normativo se centra en lo que debería ser, más que en lo que es; asimismo, enfatiza el nivel de resultados esperados de un determinado desempeño, conduciendo por lo regular a un ejercicio comparativo entre lo obtenido y un referente, norma o nivel aspirado. La información acerca del impacto revelará si las condiciones o eventos observados tienen alguna relación o pueden ser atribuidas a la operación de la política (producto de la intervención).
4. Recolección de evidencia. El equipo evaluador se encargará de usar las técnicas y los medios recomendados para el acopio de información. Esta actividad puede asociarse al trabajo de campo de cualquier investigación, desplegando recursos logísticos

para acopiar la evidencia necesaria y, por tanto, determinará el uso de información primaria o secundaria, según sea el caso. En el primero de los casos identificará si la información utilizará encuestas, entrevista, grupos focales, etc., cada uno de los cuales se asocia con medios de recolección de datos, a partir de la observación, la interacción con actores o la documentación de actividades. Para fuentes secundarias, verificará la pertinencia, disponibilidad y validez de la información que se presume servirá como evidencia. Siempre se debe tener presente que el registro de la información constituye un momento central, pues constituirá la base para futuras recomendaciones y emprendimientos en torno a la actividad realizada.

5. **Análisis de resultados.** Bajo este se emprenderán actividades correspondientes a la verificación, codificación y tabulación de datos, análisis, estratificación o categorización y comparación. Si bien, se trata de un ejercicio de gabinete, esta actividad demanda el mayor rigor y precisión, por lo que el uso de la tecnología y la validación de los hallazgos, entre otros, garantizarán la calidad del ejercicio analítico y permitirán identificar los elementos centrales para la redacción del reporte final. La disposición de herramientas o software especializado, en buena parte determinará la calidad de los resultados a reportar. El equipo debe garantizar las habilidades requeridas para emprender el análisis.
6. **Reporte de hallazgos.** La difusión de los hallazgos es una condición necesaria en todo ejercicio evaluativo. Por ello, no debe olvidarse que la evaluación, en esencia, es una actividad prescriptiva y, en consecuencia, se esperarían recomendaciones. El reporte, por tanto, debe contener insumos (sugerencias, recomendaciones, observaciones, etc.) para la toma de decisiones de los responsables del programa y para consideración de los actores interesados en el programa o política. Debe ser un documento accesible, claro y conciso que pueda ser comprendido, validado y asumido por cada uno de los interesados en los hallazgos de la evaluación.

3.3 REFLEXIÓN ADICIONAL ACERCA DE LAS EVIDENCIAS EN LA EVALUACIÓN

Por ser una cuestión crucial para los fines de este documento, se realiza una reflexión adicional al diseño de la evaluación y a la recolección de evidencia necesaria, intentando dar mayores elementos para quienes deciden emprender dicha actividad. Se reconoce que estos momentos constituyen los mayores retos y, en consecuencia, se asume necesitar mucha ayuda y soporte. Inicialmente, se propone que la clave del éxito en toda evaluación está en formular interrogantes de forma concreta y, por supuesto, responderlos de manera rigurosa, garantizando así la utilidad del ejercicio. De igual modo, el equipo evaluador debe:

- a) Establecer la clase de información a acopiar. Dado que la información es un recurso escaso y costoso, debe establecerse con antelación todo lo relacionado, definiendo, por ejemplo, si es primaria o secundaria, si es cuantitativa o cualitativa. Dicha aclaración afecta los costos e impone restricciones a su precisión; de paso, exige habilidades para diseñar instrumentos, emprender pruebas piloto y establecer contacto con las fuentes primarias. Siempre se exige verificar la validez de la información, toda vez que es un determinante de la calidad de los resultados.
- b) Indagar sobre las fuentes de información. A partir de la naturaleza de la información, establecer la(s) fuente(s) de información para ponderar su confiabilidad. Dado que la fuente afecta la calidad de la evidencia, debe procurarse su veracidad, ya que las recomendaciones estarán basadas en la información que genere. Conviene evitar sesgos u omisiones tendenciosas e innecesarias y por razones de economía frente a información primaria; en caso de ser necesario, deben considerarse los métodos de muestreo. Y es que, por razones de economía y limitaciones logísticas, no siempre será posible consultar a toda la población atendida o impactada por una política. En consecuencia, debe determinarse la estrategia que permitirá recolectar la información requerida por medio de las distintas opciones de muestreo de las que se disponen. No son decisiones menores, toda vez que lo acordado tiene efectos en la calidad de la información y en los costos de su acopio.
- c) Determinar la oportunidad y frecuencia de recolección de datos. Responder interrogantes en torno a quién, cuándo, dónde y cómo se aplicará el instrumento y recolectará la información, orientará las decisiones cruciales para un despliegue logístico eficiente y efectivo. Esto refuerza la idea de que el diseño de una evaluación requiere una planeación detallada y que nada puede dejarse al azar, so pena de enfrentar dificultades durante su implementación. En ocasiones, la estrategia comparativa proporciona insumos importantes para tomar decisiones, pero el ejercicio comparativo tiene sus reglas y limitaciones. En efecto, deben establecerse los distintos criterios para su realización y no siempre se está en capacidad de emprenderlo o, simplemente, carece de sentido. El rigor investigativo impone así condiciones que el equipo evaluador debe tramitar con antelación.
- d) Decidir la estrategia de análisis de información. Anticipar las técnicas y métodos para el análisis de la información, ya que, en ocasiones, primero se recoge la información y luego se decide la estrategia de análisis. Así se evita el desperdicio de esfuerzos y recursos, al recopilar información irrelevante. Por ninguna razón deben escatimarse esfuerzos para obtener el mejor diseño, esto es garantía para enfrentar dificultades y procurar el uso eficiente de los recursos.

La mejor alternativa para lograr resultados notables es un diseño riguroso, evitando evaluaciones mediocres con resultados pobres y poco significativos. La pretensión debe ser alcanzar

un diseño sólido que no tenga semejanza a ningún otro, una suerte de diseño inédito y exclusivo. La valoración del diseño de la evaluación no se descarta, implica entonces revisar si los supuestos del diseño son realistas y si pueden darse por sentadas las condiciones previstas para el proceso evaluativo.

3.4 LA ESTRATEGIA EN EL DISEÑO

En general, un buen diseño de evaluación se distingue por la forma como se formulan las preguntas de evaluación, esto permite determinar la clase de información a ser acopiada, identificando fuentes, espectro, métodos, instrumentos, momento y frecuencia de la recolección y plan de análisis. También pesa la forma como se ejecutan las estrategias metodológicas para responder esas preguntas. Siempre será bueno un diseño que incorpore un plan de recolección de datos, anticipe y trate los problemas y obstáculos, y detalle el análisis, asegurando que las preguntas se respondan con los datos apropiados. El diseño de la evaluación evita incurrir en sobrecostos, tiempos muertos, información innecesaria, duplicación de actividades y esfuerzos, y, por lo común, recurrir a actividades innecesarias e inútiles que afectarán la estructura de costos y generarán confusión y malestar entre los participantes. En general, un diseño riguroso garantiza un buen uso del tiempo, la calidad del estudio y una guía de la recolección de datos. Dos advertencias con relación al diseño de cualquier evaluación son:

1. No hay recetas para el diseño, pues este emerge de las preguntas de evaluación, las restricciones imperantes y los recursos disponibles.
2. El diseño no contempla una sola estrategia y es probable que sea resultado de una combinación de estrategias de evaluación. A este tenor pesa mucho la experiencia previa y la creatividad del equipo evaluador.

3.4.1 Diseño estratégico de una evaluación

Generar un diseño es pensar estratégicamente, establecer el vínculo entre las preguntas formuladas y las formas de recolectar y analizar datos para encontrar respuestas acertadas. Por tanto, hacer la(s) pregunta(s) correcta(s) son aspectos fundamentales de cada diseño; por el contrario, el estudio puede ir en una dirección inadecuada. Cronbach (1982, en GAO, 1991) advierte que a la hora de formular preguntas de evaluación se debe:

1. Considerar un amplio espectro de preguntas potenciales, incluso si no parecen plausibles o defendibles. En ocasiones, se acostumbra a hacer preguntas de acuerdo con las etapas del programa, como: ¿cuáles son sus objetivos?, ¿qué etapas tiene su proceso?, ¿qué está pasando con el programa?, interrogantes que brindan la informa-

ción para tener un panorama detallado de lo que intenta lograrse. De igual manera, conviene indagar acerca de sus resultados e impactos, como: ¿cuáles son las consecuencias de la intervención para los beneficiarios?, ¿la acción emprendida llevó a la obtención de dichas consecuencias? Esto, para descartar que el azar u otro tipo de circunstancias hayan llevado a consecuencias que no son derivadas de la acción de la política. La causalidad, de igual forma, debe verificarse; esto es, establecer el vínculo entre los procesos y resultados concretos, por ejemplo, ¿el programa de entrenamiento posibilitó las habilidades que el beneficiario demuestra? Los niveles explicativos se logran respondiendo: ¿qué pasó?, ¿cómo pasó?, ¿por qué pasó?

2. Contrastar las preguntas con los recursos disponibles para el ejercicio de evaluación. Decidiendo así preguntas factibles de responder sin caer en ilusiones ni generar falsas expectativas. El evaluador debe ser consciente de las restricciones imperantes para que la evaluación genere un diseño con la consonancia y la calidad requeridas. Para decidir sobre las preguntas es necesario tener en cuenta: el cronograma decisional, el poder relativo de las partes interesadas, las preferencias de los grupos de interés, las incertidumbres en el conocimiento base, la practicidad y los supuestos de la teoría del programa, el uso potencial de las conclusiones o hallazgos y el juicio profesional del evaluador.

Las preguntas bien formuladas privilegian las oportunidades de respuesta, pero si se pregunta demasiado a los usuarios y al personal, se puede perder apoyo y colaboración. De igual manera, al preguntar en exceso se elevan las demandas de información a recopilar, codificar, digitalizar y analizar, corriendo el riesgo de una desactualización. Se recomienda también la selectividad y la racionalidad a la hora de preguntar, además de:

- Reunir la información necesaria. La regla de oro es una correspondencia entre la evaluación y la naturaleza del programa. Una evaluación a gran escala y costosa se justifica si el programa es amplio y de largo aliento.
- Determinar las restricciones de la evaluación. Siendo el tiempo la limitación más importante, es preciso una planeación en tiempo real. Los costos son otra limitante y determinante de los métodos a utilizar, la clase de información, la tecnología, etc.; sin embargo, la calidad de la evaluación no debe depender de los costos. La experiencia del personal, en ciertas oportunidades, es una restricción; por tanto, la evaluación no debe ser más compleja que las habilidades del equipo. La ubicación de las dependencias, en determinado momento, puede actuar como restricción secundaria.
- Seleccionar las estrategias. Dependiendo de la clase de estudio a emprender, se tratará de encuestas a una muestra, estudios de caso, experimentos de campo o uso de datos secundarios. Cada estrategia encarna un diseño particular a elección del evaluador.

Diseñar un plan de evaluación toma tiempo y exige destrezas, así como un especial cuidado que debe otorgarse a la formulación de las preguntas y a los medios para obtener las respuestas. Al inicio toma tiempo, pero posponerlo o eliminarlo equivale a incursionar en demoras costosas, recolección de datos incompleta o mediocre y análisis inciertos.

3.4.2 Tipos de diseños de evaluación

Si bien, son muchas las tipologías propuestas, a continuación se hará referencia a las encuestas muestrales, los casos de estudio y los experimentos de campo (Runkel y McGrath, 1972; Judd y Kidder, 1986; Black y Champion, 1976 en GAO, 1991).

- a) Mediante encuestas se recolectan datos de una muestra representativa de la población, a fin de determinar aspectos propios de la política como incidencia, procesos, beneficios, etc. La idea es acopiar evidencia de una muestra para hacer la respectiva inferencia sobre la población. El muestreo, por tanto, debe ser riguroso para garantizar la representatividad y facilitar la generalización. Cada elemento de la población debe tener la misma probabilidad de ser seleccionado en el diseño muestral. El instrumento de recolección apoyará entrevistas personales, diligenciamiento de cuestionarios o encuestas cara a cara, telefónicas o por correo electrónico. La unidad de análisis puede estar encarnada por individuos, hogares, instituciones o territorios de cobertura de la política. En este tipo de ejercicios, las preguntas descriptivas y normativas prevalecen y, en consecuencia, cierta cautela debe observarse sobre los juicios emitidos. Lo presentado hasta aquí no difiere en mucho de lo que pudiera hacerse al diseñar una investigación en sus asuntos operativos.
- b) Los estudios de caso. Descripción analítica de un evento, proceso o programa/política (Hoaglin *et al.*, 1982 en GAO, 1990). La información recolectada debe garantizar cabal entendimiento sobre una situación compleja por naturaleza, los procesos, los productos y los propósitos, a fin de establecer los patrones subyacentes que soportarían las explicaciones. Eso indica que las preguntas de evaluación poseen un doble propósito al ser de carácter descriptivo y de impacto. Asimismo, los casos pueden ser simples o múltiples y aunque la muestra suele ser pequeña, la información debe ser densa y profunda. Según el rigor del ejercicio, algunas hipótesis pueden proponerse para determinar la causalidad existente y lograr lo que se denomina la “inducción analítica”. El caso más simple considera una única unidad, proceso, entidad en torno a un programa o política, por lo que puede ser exploratorio para conocer o ilustrar las particularidades de un evento, sin pretender hacer generalizaciones ni teorías conclusivas.

- c) Los experimentos de campo. Establecen el impacto de una intervención política; por tanto, se hacen inferencias causales acerca de los impactos generados por el programa o política bajo estudio. En el marco de experimentos tradicionales, el evaluador compara dos grupos, uno que ha sido beneficiado por el programa y otro no beneficiario. A la luz de una pregunta general, intenta establecer las diferencias entre estos grupos; diferencias, desde luego, atribuibles al programa. Tres tipos de experimentos más nombrados son:
- I. real, al elegir unidades de forma aleatoria para un grupo de tratamiento y para otro de control, a fin de compararlos;
 - II. comparación no-equivalente entre grupos, con la intención de establecer el impacto originado por la intervención; sin embargo, la asignación de unidades a los grupos de tratamiento y control no es aleatoria;
 - III. la perspectiva del antes-después, en donde no se conforman grupos de comparación, sino que se estudian las unidades antes de ser intervenidas y después de la intervención. Se supone que existen cambios en estos dos momentos y estos deben ser atribuibles a la política a la que fueron sometidas. De esta manera, podría verificarse el impacto derivado de haber estado bajo la acción del programa o política.

Aunque, en nuestro medio, la evaluación experimental no es muy usual, es una posibilidad que otorga ventajas en rigor y magnitud de la actividad. La única advertencia es que los asuntos éticos deben resguardarse, ya que, en ocasiones, constituyen la mayor restricción para emprender experimentos.

3.5 VALORACIÓN DEL DISEÑO DE LA EVALUACIÓN

La División de Metodologías y Evaluación de Programas de la *General Accounting Office* de los Estados Unidos desarrolló un sistema de revisión sistemática de la fase de diseño para responder al interrogante: ¿vale la pena dedicar tanto tiempo a la planeación y a la evaluación del diseño? (GAO, 1991). Al respecto, el evaluador debe entender lo que ha diseñado para controlar su ejecución, pero, además, tener claro lo decidido y lo omitido, así como las fortalezas y debilidades de su diseño, las necesidades del cliente y la contribución del diseño para su satisfacción. También, conviene determinar si las restricciones (tiempo, costo, personal, etc.) han sido tratadas de forma adecuada. Para poner a prueba el diseño de la evaluación se sugiere dar respuesta a las siguientes preguntas (GAO, 1991):

1. ¿Es completo? ¿Cuáles son sus limitaciones?
2. ¿Qué tan factible es su implementación en el tiempo y con los recursos disponibles?
3. ¿Está acorde con las necesidades de información, conclusividad y oportunidad?
4. ¿Hay coincidencia entre los intereses, las perspectivas teóricas y prácticas entre el contratante y el evaluador?

La respuesta a estos interrogantes privilegiará la rigurosidad del equipo en cuanto al planteamiento evaluativo, anticipando necesidades y recursos, determinando la oportunidad de estas necesidades y recursos, y evitando incoherencias metodológicas y teóricas. Lo más importante es que ofrecen la seguridad, confianza y certeza acerca de la tarea a emprender.

3.6 LOS INDICADORES: INSTRUMENTOS PARA LA EVALUACIÓN

Aunque después se volverá sobre el tema, es importante anticipar que la evaluación saca grandes beneficios del uso de los indicadores y que estos nacen en el preciso momento del diseño de una evaluación. Un indicador, por tanto, es un instrumento generado de un proceso evaluativo que permite medir cambios en el tiempo y son, en esencia, medios para determinar el progreso o retroceso de un resultado, sean estos intermedios o finales y, por supuesto, para verificar el cumplimiento de los objetivos a largo plazo. En el sector público, por lo regular, son utilizados para evaluar atributos o asuntos del progreso y del desarrollo.

El diccionario define indicador como mecanismo que proporciona información sobre algún fenómeno (presión, temperatura, etc.), bajo la forma de un valor (estadística o *ratio*) derivado de una serie de observaciones. Si bien, los indicadores son referidos como colecciones de estadísticas, suelen asociarse como números simples (a menudo, una *ratio* o, algunas veces, un compuesto de más de una) que intentan capturar una realidad más o menos compleja; por ejemplo, el ingreso per cápita, el crecimiento promedio del empleo o el promedio de grados escolares de los jefes de familia (De Vries, 2011). Un indicador “es por definición una simplificación de la realidad, algunas veces un constructo real [...] un simple número, razón o proporción u otro hecho observado que sirve para evaluar una situación” (De Vries, 2011, p. 319). El indicador es una expresión cualitativa o cuantitativa observable que describe características, comportamientos o fenómenos de la realidad mediante la evolución de una variable o del establecimiento de una relación entre variables que, comparada con periodos anteriores, productos similares, metas o compromisos, permite evaluar el desempeño y su evolución en el tiempo (DANE, 2009).

En el sector educativo, los indicadores pueden concebirse de distinta manera. Se centran más en efectividad y eficiencia, que en otras cuestiones como la justicia social y la democracia; describen y prescriben, pero no explican ni son críticos (Barnetson y Cutright, 2000).

Los indicadores de desempeño son instrumentos de política (*policy instruments*), ya que motivan a individuos e instituciones a actuar (Barnetson y Cutright, 2000). Vistos así, los indicadores son herramientas de política diseñados para promover una agenda, al combinar enfoques basados en prioridades e incentivos para implementar políticas, y formulados para generar un determinado conjunto de resultados. Bajo esta mirada, el desempeño institucional es un foco de atención, pero también una expresión de la forma en cómo se ejerce el poder, en la medida en que los indicadores reflejarían propósitos, valores o metas de un gobierno y lo que se esperaba de sus agencias. Como instrumentos de política facilitan el uso de incentivos y castigos financieros que manipulan la conducta institucional.

Cualquier conjunto de indicadores tiene a la vez, componentes normativos y selectivos acerca de la forma como deben concebirse los asuntos, estructurando sistemas para la toma de decisiones. En consecuencia, reducen la autonomía y la discrecionalidad de los responsables, otorgando unidad a los propósitos institucionales.

3.6.1 Indicadores sociales

La pretensión de las iniciativas de indicadores, como todo sistema de información, descansa en la necesidad de monitorear el desempeño gubernamental y supranacional. El uso de cualquier sistema de información tiene sustento en la noción de la responsabilidad (*accountability*), bien de los gobiernos o, en su defecto, de las agencias que los representan. En este sentido, subyace el principio que cuando el capital de algunos se usa para financiar emprendimientos, se debe estar en capacidad no solo de juzgar el desempeño de quienes actúan en su representación, sino también de imponer sanciones. En otras palabras, un buen sistema de información debería permitir, al *principal*, ejercer un control satisfactorio sobre su *agente* (Smith, 1989).

La Organización de las Naciones Unidas (ONU) hace énfasis en la cooperación para atender problemas de carácter, económico, social, cultural o humanitario; estos propósitos son monitoreados por medio de indicadores construidos para determinar su logro. El uso de indicadores surgió con el inicio de la reconstrucción de los países tras la Segunda Guerra Mundial y el artículo 55 de la Carta de las Naciones Unidas (1945) establecieron sus responsabilidades para la promoción de mejores niveles de vida, progreso y desarrollo económico y social. En consonancia, un grupo de expertos de esta organización publicó, en 1954, un primer documento acerca de la definición y medición internacional de niveles de vida. Desde entonces existen versiones actualizadas y mejoradas acerca de las directrices, metodologías e indicadores para medirlos, identificando los factores económicos y sociales determinantes. Los indicadores son:

[S]eries estadísticas y toda forma de evidencia [...] que nos permite evaluar dónde estamos y hacia dónde vamos respecto a valores y metas, y evaluar programas específicos y determinar su im-

pacto. [...] Representan conceptos complejos o condiciones, son estadísticas u otras formas de evidencia que intentan dar sentido a lo incierto o lo desconocido, extrayendo ideas simples de las complejas (United Nations, 1999, p. 27).

“[U]n indicador es una estadística o un conjunto de estadísticas, u otra clase de evidencia apropiada para evaluar una situación o un desarrollo, con relación a cierta meta acordada” (United Nations, 1999, p. 26). El Índice de Desarrollo Humano (IDH) propuesto por el Programa de las Naciones Unidas para el Desarrollo (PNUD), por ejemplo, intenta capturar los niveles de bienestar de las comunidades a partir de tres aspectos: vida larga y saludable, educación y nivel de vida digno. Cada asunto encarna dificultades para su valoración, medición e integración. A pesar de estas dificultades, un indicador sintético (cuantitativo que va de 0 a 100, cero representa el menor estado de desarrollo), refleja los niveles de bienestar y desarrollo humano, y permite la comparación de las actuaciones gubernamentales.

La intención de cualquier índice es captar algo de suyo difícil, como lo es el desarrollo en sus dimensiones más esenciales. Como cualquier indicador presenta limitaciones e imperfecciones, pero ha sido una manera sencilla de establecer los niveles en cada área geográfica y, algo adicional, realizar las comparaciones en los últimos años. También ha servido para enfatizar las acciones gubernamentales en sus metas de desarrollo y de ahí su importancia o relevancia para las políticas públicas sociales.

3.6.2 Atributos de los indicadores

Dada la importancia de estos instrumentos para emprender los juicios del actuar social y gubernamental, se esperaría que tuvieran un conjunto de atributos para el proceso de toma de decisiones. A lo largo de los estudios de la evaluación y el desarrollo, los estudiosos han reflexionado y encontrado algunas características a ser tenidas en cuenta. En primera instancia, un indicador debe ser válido y confiable; esto es, debe medir lo que pretende medir de forma sistemática y evitando la subjetividad. En segundo lugar, deben ser comprensibles, con un significado evidente e inequívoco para todos. También deben ser oportunos, calculados con cierto grado de certeza para los tomadores de decisiones; de otra manera, deben ser útiles para dirigentes y publi-administradores. Otro atributo de gran relevancia es la resistencia a comportamientos indeseables, es decir, no generar incentivos inadecuados ni conductas reprochables. A este tenor, es evidente que algunas personas, en su afán por cumplir con los indicadores, subvierten asuntos legales o morales.

Los indicadores, en complemento, deben ser integrales al capturar las dimensiones más importantes del actuar de las personas y las organizaciones. Adicionalmente, no deben ser redundantes, pues cada uno aportará a asuntos distintos del desempeño bajo los principios de la economía. Asimismo, deben ser comparables en el tiempo y el espacio, así como entre las

personas, las organizaciones y las instituciones. Esto permite establecer medidas correctivas e instaurar planes de mejora continua. Por último, los indicadores deben ser aceptables para los *stakeholders* o grupos interesados en los resultados, dicha aceptación expresa su consentimiento. Visto así, la construcción de un indicador encarna desafíos importantes y demanda el máximo rigor científico posible.

3.7 SÍNTESIS CAPITULAR

A la luz de los desarrollos hasta aquí presentados, podría afirmarse que la evaluación es una actividad, en esencia premeditada; esto es, exige con anterioridad un proceso sistemático y que, en consecuencia, obliga a que sus responsables hagan los mejores esfuerzos para mantener el control de su dinámica. Este capítulo ha puesto de relieve tal premisa y ha señalado los elementos integrantes del proceso que rigen una actividad evaluativa de manera detallada para que los encargados tengan una guía, como también un control sobre su desarrollo de principio a fin. La consigna y recomendación es que se tenga el control de todo el proceso y que no se deje nada al azar ni asuntos a resolver sobre la marcha.

También, como pudo apreciarse, existen muchas alternativas o posibilidades para el diseño de una evaluación y el equipo responsable tiene que escoger la más adecuada y oportuna. De un buen diseño depende, en buena parte, el éxito final y vale la pena hacer los mejores esfuerzos para determinar con anticipación todas las etapas, técnicas y medidas necesarias para conceptuar sobre los atributos de la política o programa. Ello es y será siempre un imperativo.

REFERENCIAS

- De Vries, W. (2001). Meaningful measures: Indicators on Progress, Progress on Indicators. *International Statistical Review/Revue Internationale de Statistique*, 69(2), 313-331.
- Glenaffric, Ltd. (2007). *Six Steps to Effective Evaluation: A handbook for programme and project managers*. The Joint Information Systems Committee. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.436.3108&rep=rep1&type=pdf>
- Larkin, S. (ed.). (2001). *Pautas para Evaluaciones. Pasos para Organizar una Evaluación*. Tearfund.
- Meyer, W. (2015). *Introduction to Evaluation Design*. [http://new.ceval.de/modx/uploads/pdf/Evaluation%20Design%20-%20Block%20%20A%20-%20Introduction%20\(10-07-15\).pdf](http://new.ceval.de/modx/uploads/pdf/Evaluation%20Design%20-%20Block%20%20A%20-%20Introduction%20(10-07-15).pdf)
- Nash, R.; Hudson, A. y Luttrell, C. (2006). *Mapping Political Context: A Toolkit for Civil Society Organisations*. Overseas Development Institute. <https://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/186.pdf>
- National Science Foundation (2002). *The 2002 User-friendly Handbook of Project Evaluation*. <https://www.nsf.gov/pubs/2002/nsf02057/nsf02057.pdf>

- República de Colombia. Departamento Administrativo Nacional de Estadística, DANE (2009). *Guía para diseño, construcción e interpretación de indicadores*. https://www.dane.gov.co/files/planificacion/fortalecimiento/cuadernillo/Guia_construccion_interpretacion_indicadores.pdf
- Smith, P. (1990). The Use of Performance Indicators in the Public Sector. *Journal of the Royal Statistical Society*. Series A (statistic in Society), 153(1), 53-72.
- United Nations (1999). *Integrated and coordinated implementation a follow-up of major United Nations Conferences and Summits*. A critical review of the development of indicators in the context of conference follow-up. <https://digitallibrary.un.org/record/272664?ln=es>
- United Nations. United Nations Development Program, UNDP (2009). *Handbook on planning, monitoring and evaluating for development results*. <http://web.undp.org/evaluation/handbook/documents/english/pme-handbook.pdf>
- United States. General Accounting Office (1990). *Case Study Evaluations*. Washington: Program Evaluation and Methodology Division. GAO/PEMD-91-10.1.9.
- United States. General Accounting Office (1991). *Designing Evaluations*. Washington: Program Evaluation and Methodology Division. GAO/PEMD-10.1.4.
- Wall, J. (2004). *Program Evaluation Model. 9-Step Process*. http://www.janetwall.net/attachments/File/9_Step_Evaluation_Model_Paper.pdf

CAPÍTULO 4

Ética y evaluación

INTRODUCCIÓN

A pesar del permanente reclamo sobre los asuntos éticos del sector gubernamental, los escándalos crecen, en ocasiones ante la indiferencia de la autoridad o, incluso, con su bendición. La evaluación como práctica profesional o elemento insustituible en el desempeño del sector público postula que, su distancia con la ética es corta, pero, aun así, exige un mayor apego a sus imperativos. En particular, el rol del evaluador como vector ético es indiscutible, por lo que reflexionar sobre el vínculo entre ética y evaluación resulta pertinente.

En la evaluación, la ética se refiere a un conjunto de estándares normativos para ayudar a resolver los dilemas que, a menudo, enfrentan los evaluadores. Más allá de los reportes de evaluación, los beneficiados de esta observancia ética son el bien común y el desempeño político. De otra manera, estaría en cuestión la legitimidad de la práctica, el uso de los recursos, el bienestar y el interés de las partes, asuntos que no se deben poner en riesgo.

Este capítulo tiene como propósito poner de relieve aquellos postulados y lineamientos necesarios para lograr que las prácticas evaluativas observen y, en ocasiones, recuperen la pulcritud y transparencia demandadas. Esto es, que la práctica evaluativa incorpore, sin cortapisas, las consideraciones de la ética aplicada. De ninguna manera puede considerarse un estudio exhaustivo sobre los asuntos éticos, los cuales, por su extensión y especialización, rebasarían los propósitos y competencias de esta obra. Simplemente, intenta resaltar la importancia de la ética a la hora de emprender una evaluación y de preservar sus lineamientos a la hora de juzgar los atributos de una política pública. Asimismo, se parte de una breve consideración teórica de la ética, después se recupera y reflexiona en torno al rol profesional del evaluador. Sigue con la exposición de las características del contexto en que toma lugar la evaluación, el ambiente político de los *stakeholders* y las condiciones que privilegian la ética en la evaluación. A continuación, se reseña el papel de la ética en las distintas fases de la evaluación para extraer las implicaciones de un concepto o juicio evaluativo y termina discutiendo la necesidad de regular éticamente el ejercicio evaluativo.

4.1 GENERALIDADES ACERCA DE LA ÉTICA

Podría asegurarse que, en la actualidad, la ética se niega a desaparecer y que, a pesar de las permanentes contravenciones a sus postulados, el reclamo social es creciente y una necesidad urgente. Indudable que en la actualidad esta rama de la filosofía enfrenta desafíos teóricos y prácticos importantes al punto que, en ocasiones, no soportan la comprensión de razones que motivan elecciones inmorales. Esta disciplina trata con la moralidad y los sistemas morales y se entiende como la “sistematización de las reglas y normas relativas a la moralidad tal y como se practica, regulando el comportamiento dentro de una determinada comunidad” (Breczko y Breczko, 2018, pp. 5-9). De manera general, aborda los difíciles asuntos relativos al bien, al deber, a la conciencia, a la responsabilidad y sobre lo que deben ser las metas o propósitos de la acción humana. Desde los postulados en la antigua Grecia intenta establecer algo sencillo, pero de difícil resolución: cómo vivir bien, plácidamente y cómo alcanzar la felicidad individual y colectiva.

Como campo de estudio, el desarrollo teórico de la filosofía moral es antiguo, denso y, en ocasiones, hasta difuso. A pesar de su importancia, no es sencillo de abordar, ni fácil de difundir o aplicar. Existen falsas premisas o creencias convenientes, de alguna manera, para reducir la ética a las opiniones o la intuición de los individuos. En ocasiones, también se considera un asunto reservado para seres virtuosos, olvidando que sus principios tienen un entronque significativo con la toma de decisiones (una de las áreas de mayor intensidad humana). Por ello, la filosofía moral reclama una dimensión aplicada y cada día toma más fuerza su incidencia para orientar las decisiones y las acciones de los seres humanos (Cortina, 2005). Los preceptos éticos y su aplicación en la vida cotidiana solo ofrecen un marco reflexivo para la toma concreta de decisiones. No se aconseja una aplicación de principios generales a casos concretos, como tampoco derivar máximas de decisiones concretas, más bien “descubrir en los distintos ámbitos la peculiar modulación del principio común” (Cortina, 2005, p. 16). Todo esto tiene que ver con el “desempeño adecuado de la función evaluativa en el sector público”, quizá uno de los últimos bastiones de la acción colectiva para recuperar su dignidad.

La ética profesional o deontología profesional (del evaluador) derivó de la noción del deber moral (obligación moral). Sus estándares están definidos en un conjunto de normas deontológicas o códigos profesionales establecidos por profesiones liberales o de confianza pública. Estas consideraciones, sin duda, constituyen un referente importante para el rol del evaluador cuyo actuar está alinderado por las cuestiones éticas y condicionado por las presiones y consecuencias derivadas de sus juicios. La evaluación tiene muchas particularidades en su ejercicio y quienes la ejercen deben ser conscientes de las circunstancias a enfrentar, de los dilemas a resolver y, sobre todo, dilucidar cómo su actuar y su rol se pueden poner a salvo en circunstancias riesgosas. Es importante reconocer que la función no es sencilla y que su desempeño exige competencias y claridades. Pero, ¿cuál es el rol del evaluador?, ¿cuáles son sus

vínculos con los principios éticos? A continuación, una sección trata las cuestiones propias de la acción investigativa que deben guiar su misión en la generación de conocimiento. Después, se enfatiza su misión educativa y, por último, el énfasis se hace en su rol como representante de causas supremas, como el bien público y el interés general.

4.2 EL ROL DEL EVALUADOR

Como fue postulado, la función evaluativa reconoce unas premisas y condiciones que vale la pena reflexionar. Varios han sido los aportes de autores acerca de las responsabilidades conexas a la labor del evaluador. El evaluador y su comportamiento están guiados por un patrimonio axiológico propio, acumulado, transformado y consolidado, que configura un recurso importante para emitir juicios, más allá de otros condicionantes que pudieran intervenir cuando se evalúa. Y es que juzgar requiere sabiduría, juicio y justicia, so pena de generar daños innecesarios.

Reconocidos teóricos de la evaluación (Scriver, 1969; Patton, 1997; Weiss, 1998, y Ryan y Schwandt, 2020) han abordado el papel de evaluador. Desde el análisis de las políticas, al analista se le han otorgado responsabilidades de evaluador, de ahí que su papel sea múltiple y apenas lógico, ya que evaluar encarna funciones de carácter científico, administrativo, social y político. Evalúa, entonces, capacidades y habilidades particulares y, a la vez, exigentes, de acuerdo con los autores mencionados; a continuación, se describen sus roles representativos.

4.2.1 Experto

Sin ánimo de magnificar su función, es de reconocer que el evaluador es un generador de conocimiento en los distintos momentos del proceso de juzgamiento. Emprende diversas actividades científicas con el ánimo de arribar a un juicio; por sus antecedentes académicos y laborales es una voz autorizada para conceptuar. El experto en cualquier asunto (salud, educación, cultura, etc.) posee habilidades o conocimientos al respecto. En ocasiones, existe un respaldo formal de instituciones que avalan dicha experticia que lo habilitan para realizar actividades evaluativas que van desde el diseño de la intervención hasta la emisión y difusión de los hallazgos. Usa conocimientos de la investigación social, no solo para desplegar una estrategia metodológica para la generación de conocimiento en torno a una situación particular, sino también para conceptuar sobre el desempeño alcanzado.

No se descarta que la experticia del evaluador esté modulada por las diversas tradiciones epistemológicas de su formación. Por todos es sabido que estas tradiciones, en algún momento, podrían incidir en el actuar evaluativo. Es pertinente, por tanto, distinguir el perfil evaluativo del positivista, del interpretativista, del crítico y del experimental. El positivista preferirá datos secundarios, trabajo de gabinete, técnicas estadísticas para el análisis; todo es-

to le permitirá un juicio objetivo. Si la preferencia del evaluador es interpretativa, hará trabajo de campo, diseñará instrumentos de recolección de información, recogerá testimonios y hablará con todas las partes involucradas e interpretará sus hallazgos. Una visión crítica del evaluador le indicaría estar del lado de los beneficiarios, denunciando cualquier anomalía de la política y sugiriendo medidas de fondo para alcanzar objetivos más ambiciosos y trascendentales de la intervención. Lo claro es que la corriente epistemológica marca ciertas preferencias, actitudes e ideales a la hora de investigar, emitir juicios y sugerir correctivos.

Las competencias del experto le orientarán para definir, de manera acertada, los procedimientos, instrumentos y, en general, las estrategias de evaluación e investigación que conduzcan a la generación de conocimiento que alimente cualquier proceso decisional de quienes poseen dichas competencias. El rigor científico podrá ser constatado en el diseño de la evaluación, el cual es una función del experto en evaluación, mismo que lo autoriza para guiar acertadamente los equipos de investigación y generar el ambiente propicio para las actividades evaluativas. Este respaldo le permitirá generar las sinergias y ambientes adecuados para administrar los intereses disímiles que reinan en estos contextos. “La mirada independiente del experto sobre el trabajo propio genera ideas que ayudan a mejorarlo” (Scriver, 2009, iv).

El experto, de igual manera, hará patente sus virtudes científicas en las actividades para la elección de los instrumentos adecuados y suficientes, la recolección de datos y análisis de resultados, el procesamiento acertado de datos, y la administración y custodia de la información. La economía de la información será uno de sus logros, así como la organización, representación y difusión acertadas. Estas funciones hacen la diferencia entre expertos y quienes no lo son, basta con asistir a una de sus reuniones o revisar uno de sus documentos para ponderar su experticia.

Más allá de la administración adecuada del proceso de intervención, los resultados emergen y deben ser juzgados. Este proceso de ponderación no se hace en el vacío, se apoya en la evidencia, pero también en la experticia y en los valores rectores. Se aprecia que el razonamiento es una de las funciones propias de la evaluación; es la actividad que, junto a los valores, permitirían al evaluador un proceso abstracto de pensamiento para proponer juicios, identificar estrategias, acciones y maneras acordes con la concepción de su ser y criterio profesional. Dicho razonamiento, además, le permite establecer si sus conclusiones están basadas en lo correcto y están lógicamente fundamentadas. Al conjuntar el proceso de razonamiento y los valores se configura el razonamiento moral, por lo que la ética, entonces, se encuentra con la evaluación.

4.2.2 Educador

El evaluador como educador es una posibilidad interesante que presenta distintas expresiones. El juicio del evaluador informa, aconseja, asiste y facilita en lo concerniente a la toma

de decisiones y planes de mejoramiento. La función educativa de la evaluación es una de sus mayores virtudes; de hecho, en el aprendizaje la evaluación es primordial. Evaluar significa verificar si el proceso de enseñanza y aprendizaje se ha logrado en sus intenciones prefijadas. Es un proceso basado en las evidencias acumuladas de la acción-respuesta del educando y de otras variables intervinientes; es allí que, mediante el acopio de información, se enjuicia el avance y progreso de manera individual.

En consecuencia, la evaluación de una política pública se asimila a las labores corrientes de un educador, solo que las agencias públicas y sus desempeños, así como las organizaciones sociales y los mismos beneficiarios del programa son juzgados. Las dimensiones señaladas sobrepasan los elementos y tamaños tradicionales de la evaluación en el salón de clase, pero, en esencia, persiguen muchas cosas en común. Más allá de emitir juicios, la función educativa se despliega en momentos posteriores. En efecto, el juicio genera un conjunto de acciones para enmendar y mejorar los desempeños, los asuntos mediados por el consejo, la asistencia y la facilitación, derivados de la experticia del evaluador.

La evaluación, en consecuencia, no debe considerarse como aquel enjuiciamiento aislado y discrecional, cargado de severidad o dureza. Es, ante todo, una función que invita a la enmienda y al mejoramiento del quehacer de la función pública y de todo lo que implica. Compromete acciones tendientes a eliminar elementos distorsionantes, condiciones inapropiadas o actitudes inconvenientes; asimismo, encarna planes de mejora, nuevas acciones y seguimientos y apoyos administrativos y, en algunos casos, la asistencia y soporte del experto evaluador. El evaluador competente sabe que su juicio estará vinculado con futuras acciones que ayudarán a diseñar, implementar y, seguramente, evaluar; de ahí su conexión con su espíritu educador.

4.2.3 Representante

El evaluador es, sin temor a la equivocación, un representante (*advocate*), pero qué puede representar. Al parecer, lo más básico es la representación de una causa, por lo que las políticas siempre tienen causas y sendos representantes, y uno de ellos podría ser el profesional que conceptúa sobre el uso de los recursos, la función de las organizaciones o los resultados finales o intermedios. Los representantes entienden sus actividades, en términos prácticos, de cómo utilizar mejor las herramientas para alcanzar las metas deseadas de la política (Gardner y Brindis, 2017). En consecuencia, esta función de representación indica y obliga a que el evaluador trate de incidir en el proceso de las políticas, “navegando” en la arena política para proponer, argumentar y vehicular algunas acciones, estrategias, mejoras y todo aquello que sea necesario para asegurar su éxito.

En otras palabras, el evaluador termina siendo un actor político que busca incidir en las decisiones. Para ello, dispone de un conjunto de habilidades propias de la acción política, ya que moviliza, persuade y manipula para cambiar cursos de acción, pareceres o posturas ideológicas,

con el fin de mejorar y alcanzar los objetivos (pre)configurados en la agenda o en los resultados de las políticas. De manera simple, esta es la función de la representación que, dicho sea de paso, pudiera estar influida por las distintas visiones epistemológicas del agente de evaluación, como son: la objetividad, la subjetividad o la liberación por medio de la acción colectiva.

Por tanto, la evaluación es un factor de incidencia en todas las dimensiones de la vida en sociedad. No es una actividad en solitario, puede hacerse con otros actores políticos del sistema democrático. Por eso, el *policy advocacy* es el proceso que permite tomar en cuenta en las decisiones de instituciones poderosas, a las organizaciones de la sociedad civil, las redes, las coaliciones y los académicos que luchan por la justicia social y económica, la sostenibilidad y la paz.

4.3 LAS OPCIONES Y LOS NODOS ÉTICOS DE LA EVALUACIÓN

De ninguna manera la ética puede considerarse como un asunto compacto y único. El discernimiento ha sido prolífico y, por supuesto, muy recursivo y útil en diferentes dominios de la vida humana. Para el analista de políticas públicas y el evaluador, los asuntos éticos están a la orden del día: ¿magnífica los atributos y resultados de una intervención pública?, ¿oculta las malas prácticas?, ¿oculta algún daño a un beneficiario? Todos estos interrogantes encarnan elecciones éticas que no pueden evitarse y que deben tramitarse, pero cómo hacerlo. Cuando los asuntos están referidos a las acciones gubernamentales y políticas, los asuntos éticos son muy particulares.

Cuando la ética se aplica a la política, se deben atender interrogantes referidos a qué fines y metas deben servir los gobiernos y qué procesos o medios deben usar para alcanzar estos fines (Sullivan y Segers, 2007). Estas autoras sugieren tres enfoques que pudieran guiar la aplicación de la ética a la política, obvio que no son las únicas, pero son destacadas en la teoría ética. Conocida también como la ética normativa, sus tres enfoques son: el consecuencialismo, la ética deontológica y la ética de la virtud.

Se comenta que el consecuencialismo es una perspectiva observada por la mayoría de los servidores públicos. Desde allí se enfatiza en los buenos resultados como la base de la evaluación de las acciones y decisiones humanas. La idea central es que lo que hace que una acción o decisión sea correcta es porque genera mejores consecuencias que cualquier otra alternativa en competencia. Se aconseja elegir aquel curso de acción que genere el mejor resultado, el mayor impacto positivo o la máxima satisfacción de las preferencias. El famoso y renombrado utilitarismo de Jeremy Bentham y John Stuart Mill es una posibilidad en la ética consecuencialista, quienes esperaban que los individuos promoviesen el logro del mayor bien para el máximo número de personas.

A pesar de sus buenas intenciones y lo elemental de su propuesta, el consecuencialismo es centro de algunas críticas o deficiencias adjudicadas. Una de ellas es que se sacrifica el in-

terés de unos pocos por el beneficio de muchos. Otra crítica es que difícilmente se puedan anticipar todas las posibles consecuencias de una decisión o de una política. Algo relacionado con lo anterior es la dificultad de calcular las consecuencias mismas o su intensidad.

Por otro lado, la ética deontológica parte de la premisa de que existen obligaciones morales o deberes que se deben cumplir, más allá de las consecuencias derivadas de dicho cumplimiento, este prima ante la bondad o fines de la acción. En este sentido, los distintos roles que asumen las personas darían un contenido específico a las obligaciones morales o deberes. Cualquiera que sea el rol ejercido, se tienen deberes que atender y la obligación moral indica que se debe hacer de la mejor forma. La perspectiva rawlsiana también es una perspectiva etiquetada como deontológica, toda vez que se basa en el deber de tratar a las personas como fines (no como medios), condición verificada cuando consienten las acciones que los afectan. Algunas versiones de las teorías de los derechos humanos son ejemplos de los enfoques deontológicos a la ética (Sullivan y Segers, 2007).

La ética de la virtud es el tercer enfoque normativo más reconocido. Según sus proponentes, el interrogante a atender es: ¿qué clase de persona debo ser? Con raíces aristotélicas asume que existen ciertos ideales que se deben perseguir por otorgar un desarrollo total a la humanidad. Esos ideales se descubren mediante reflexión en torno a las potencialidades de las personas; asimismo, las personas desarrollan virtudes y están dispuestas, de forma natural, a actuar en consonancia. Honestidad, coraje, compasión, generosidad, fidelidad, integridad, justicia y prudencia son ejemplos de virtudes que pueden ser cultivadas por las personas. Los servidores públicos se basan en estas virtudes para decidir acciones específicas en situaciones similares. El evaluador, por supuesto, tendrá que soportarse en sus virtudes para ejercer su actividad, para que le dicten los lineamientos de su juicio y para poder tramitar las correspondientes presiones provenientes de los distintos actores.

Otros enfoques acerca de la ética dan cuenta del altruismo como valor o deber moral que guía las acciones de los individuos o grupos. El amor por los otros sirve como un estándar ético que indica que lo más importante son las personas, en especial aquellas en desventaja. De igual manera, se invoca el intuicionismo como un posible enfoque representado por D.W. Ross (1877-1971) en su obra *The Right and the Good* (1930); en esta obra se introduce una lista de consideraciones relevantes éticamente. Sus representantes introducen la lista de deberes *prima facie* como referentes en favor o en contra de una determinada acción: fidelidad, reparación, gratitud, justicia, beneficencia, autosuperación y no maleficencia. Esta perspectiva es indicativa de aquellos deberes que, a primera vista, son aceptados como correctos hasta que se demuestre lo contrario. Cumplir las promesas, por ejemplo, es una regla correcta y aceptable desde la ética (Rich, 2007). Asume también que hay cuestiones intrínsecas buenas y son ciertas por naturaleza.

La casuística es otra posibilidad invocada a la hora de resolver situaciones éticas. Es la resolución de casos específicos de conciencia, deber o conducta mediante la interpretación de principios éticos o de alguna doctrina religiosa. Se trata también de un enfoque inductivo

que lleva a la aplicación de reglas absolutas y principios a la luz de preceptos religiosos. Por último, la narrativa es una posibilidad, en especial cuando se trata de educar en cuestiones éticas; asimismo, es una estrategia para inducir ciertos comportamientos considerados moralmente convenientes.

4.4 LA DIFÍCIL RELACIÓN ENTRE LA ÉTICA, LA POLÍTICA Y LAS POLÍTICAS

El contexto de la evaluación es un contexto político, por excelencia, al que debe atenerse y adecuarse el evaluador, procurando la preservación de la función y el espíritu evaluativo. Mucho se ha advertido acerca de esta condición política para reconocer las presiones que deben soportar. Ni qué decir de los dilemas éticos que este contexto puede ocasionar, aunque suene evidente, lo primero a recordar es que la política pública tiene su origen en contextos políticos. En ese sentido, tiene propósitos político-ideológicos y se desarrolla bajo la lógica política. Dicho entorno político con todo y sus presiones desafían los valores y actuar ético del evaluador. Por ello, la evaluación se entiende como “una actividad eminentemente política, políticamente articulada y siempre operará dentro de estas restricciones” (Taylor y Balloch, 2005 en Khan, 2016).

En 1973, Carol Weiss afirmó “en donde se reúnen la política y la investigación evaluativa” e hizo tres señalamientos a la evaluación:

1. Trata con políticas y, en esencia, los programas son criaturas políticas concebidas, debatidas, legalizadas y presupuestadas en ambientes y procesos políticos. En procesos posteriores, están sujetas al apoyo o a la hostilidad del juego político.
2. Se emprende para apoyar la toma de decisiones y su reporte entra en la arena política, pues aporta evidencia acerca de los resultados y compete con otros asuntos de similar peso político.
3. En su naturaleza, la evaluación asume una postura política en la problemática tratada, en la legitimidad de sus metas y estrategias e, incluso, en el rol del evaluador.

Sin lugar a duda, los argumentos presentados y algunos adicionales, evidencian el contexto político en que se desenvuelve el agente evaluador. Ignorarlo conlleva a conflictos innecesarios, dilemas éticos y, en ocasiones, frustraciones mayúsculas.

La evaluación, entonces, es “un insumo movilizador de la agenda política o al servicio de la actividad política” (Boston, Bradstock y Eng, 2010, p. 16). Compromete, no solo, acciones políticas a niveles micro y macro e institucionales, así como las relaciones sociales y las decisiones y recursos de actores que gravitan en el proceso de las políticas. No sobra advertir

que la evaluación tiene consecuencias políticas que restringen o afectan cualquier proceso decisonal. Es necesario que sus responsables establezcan relaciones con involucrados e interesados en algún aspecto de la política (*stakeholders*). En efecto, las relaciones con los *stakeholders* son medulares en la evaluación, dada su incidencia y poder en el futuro de los programas.

De ninguna manera las relaciones con los *stakeholders* son sencillas de establecer, mantener o fortalecer, ya que su naturaleza diversa, interés focal, variopinta ideológica y otros rasgos, desafían la habilidad y la sabiduría de cualquier evaluador. Aun así, deben cultivarse, pues en dichas relaciones descansa no solo el futuro del programa, sino la efectividad y el impacto del ejercicio evaluativo. A la luz de las relaciones reales y potenciales del evaluador con todos los protagonistas del contexto, se puede asegurar que debe tejerlas hacia el interior y exterior de las organizaciones. Este doble papel lo obliga a irradiar todo su conocimiento sobre el desempeño organizacional para que sus funciones ganen en efectividad y eficiencia. A la par, debe ejercer la correspondiente influencia externa para que los *stakeholders* externos (patrocinadores, financiadores, etc.) asuman, de manera adecuada, las recomendaciones emanadas del ejercicio evaluativo y emprendan todos los correctivos del caso. La clara y oportuna difusión de la información hará que el evaluador establezca las relaciones necesarias en pro del beneficio común, una labor difícil y complicada.

4.5 GARANTIZANDO LA APLICACIÓN DE LOS IMPERATIVOS

Más allá de las perspectivas y enfoques del estudio de la ética para el ejercicio de la evaluación, las principales organizaciones han elaborado códigos de conducta para ayudar a los evaluadores a enfrentar dilemas propios. A la luz de los postulados éticos, el ambiente y el contexto en donde ocurre la evaluación no dejan de ser exigentes y complicados; por ello, distintos autores han aconsejado tener un código de ética para que el evaluador preserve su legitimidad. Esta recomendación se ha materializado en códigos de conducta y en sugerencias de las asociaciones dedicadas al fomento y fortalecimiento de las actividades evaluativas. Estas guías normativas ayudan a ejercer, de manera adecuada, la profesión como evaluador o investigador.

Dado que la evaluación es un tipo específico de investigación, los asuntos éticos de la investigación social convergen y resultarían igualmente válidos para la evaluación. Siendo así, el evaluador/investigador tiene dos compromisos superiores: generar conocimiento y garantizar los derechos de las personas que intervienen. Estos principios, en la práctica podrían entrar en conflicto, ya que avanzar en el conocimiento científico debe sopesarse a la luz de los costos de la potencial pérdida de dignidad, autoestima, privacidad o libertades democráticas. En adición, ser ético no tiene compensaciones, pero no serlo y ser sorprendido trae para el infractor castigos sociales humillantes, ruinosos y hasta legales.

En consecuencia, en la observación de los asuntos éticos, el evaluador debe considerar las reglas de la investigación científica, esto es, las reglas de la generación de conocimiento y las

de los involucrados. Las reglas de oro en cuanto al conocimiento indican que el fraude no es posible. No son tolerables, bajo ninguna circunstancia, acciones como inventar, adecuar o “maquillar” evidencias o información. La regla de los involucrados tiene que ver con los asuntos del reconocimiento de las ideas y descubrimiento de los otros o de sus colegas. Por tanto, el plagio no es permitido; reconocer los créditos a terceros no es una regla de urbanidad, sino una obligación en los círculos científicos. Cuando ocurren estas, se reconocen como prácticas científicas censurables.

Preservar los derechos de los involucrados a la par, es algo inmutable para quienes llevan a cabo actividades investigativas y evaluativas. Por ninguna razón, la investigación y actividades conexas deben causar, promover o generar daño alguno a los participantes. Al respecto, la historia de la humanidad registra experiencias suficientes que ilustran, pero que también horrorizan por su inobservancia, cuando en nombre de la ciencia humillan, degradan y lastiman a otros seres vivos para generar conocimiento.

Los códigos de ética convienen para un ejercicio evaluativo decoroso. La asociación canadiense de evaluación (www.evaluationcanada.ca) se dedica al fomento de la evaluación teórica y práctica, fijando sus principios en tres dimensiones: competencia, integridad y responsabilidad. En consecuencia, el evaluador debe garantizar la provisión de un servicio basado en los métodos apropiados de investigación, con contenidos adecuados y de frontera, y procurar un mejoramiento continuo de sus actividades. Sus actuaciones deben ser transparentes y alejadas de cualquier conflicto de intereses, antes y durante el ejercicio de sus competencias. La sensibilidad por los asuntos sociales y culturales deben guiar sus relaciones con los *stakeholders* en todo momento. La rendición de cuentas obliga a que sea responsable por la provisión de información a sus clientes, a fin de apoyar la toma de decisiones. Por tanto, los reportes deben ser claros y precisos, escritos de acuerdo con las estipulaciones científicas establecidas.

Por su parte, la Asociación Americana de Evaluación (www.eval.org) establece los mismos principios guías para quienes tienen responsabilidades de evaluación de políticas públicas. Aboga por la competencia, la integralidad, la honestidad y la responsabilidad de los evaluadores. En adición, enfatiza en asuntos éticos específicos que tienen que ver con el trato justo de las diferentes perspectivas e intereses de los individuos o grupos, contemplando inclusive aquellos que, por lo regular, no son incluidos o que son opuestos. Alienta la consideración de la ética profesional, sus normas y regulaciones (consentimiento informado, confidencialidad y prevención de daños) para los participantes en la evaluación. Con relación a los grupos de interés, recomienda mantener un adecuado balance de los intereses del cliente, de los *stakeholders* y del bien común, mientras se asegura la integralidad de la evaluación. Identifica y trata las amenazas del bien común, en especial, cuando los intereses de algún actor particular entran en conflicto con los valores de una sociedad democrática, justa y equitativa. También debe promover la difusión de la información y hallazgos con acceso equitativo, asegurando la confidencialidad, el respeto y el honor de las personas. Recomienda que

el evaluador mitigue el sesgo y cualquier desbalance de poder que pueda surgir en el contexto de la evaluación.

Otra de las experiencias que vale la pena considerar es aquella referida al Programa Internacional de Capacitación para la Evaluación del Desarrollo (IEPT, por sus siglas en inglés) ([www.https://ipdet.org](https://ipdet.org)). Iniciativa que se define como un esfuerzo colaborativo de tres socios: Centro de Educación Continua de la Universidad de Bern, Suiza (ZUW), Centro para la Evaluación en Saarbrücken (Alemania) y Grupo de Evaluación Independiente del Banco Mundial (IEG). Worthen, Fitzpatrick y Sanders (2004), advierten de las fuentes de corrupción y de las falacias que a menudo emergen en el mundo de la evaluación y que, sin lugar a duda, desafían la evaluación. Distorsionar la realidad, ocultar hallazgos y prometer asuntos a sus clientes conducen a comportamientos alejados de los estándares éticos y de la rectitud profesional esperada. El manual del IPDET (2007), de igual manera, advierte sobre los peligros provenientes de los actores externos y de quienes contratan la evaluación ya que, en ocasiones, inducen de antemano los resultados de la evaluación, o bien, alteran los hallazgos a la hora de difundirlos. Estos y otros asuntos señalados ilustran muchas de las situaciones que deben considerarse férreamente cuando se emprenden actividades de carácter evaluativo y que son más comunes de lo esperado.

El IEPT (2007) recomienda al evaluador reconocer y manejar la política para controlar sus consecuencias: identificar y clasificar los “juegos políticos de la evaluación”, de acuerdo con el rol que desempeñan los involucrados en la evaluación, así como los juegos de la gente evaluada, del evaluador y de otros *stakeholders*. Con esto, se reafirma que la evaluación, aunque expuesta a dificultades y a situaciones complicadas, debe permanecer útil para administradores, donantes, participantes y ciudadanos y que, por tanto, debe ser honesta, objetiva y justa.

Se ha sugerido que la evaluación como profesión está supeditada a los mandatos de las profesiones liberales y, en consecuencia, alentaría y atendería todas las implicaciones derivadas. Doctrinariamente, se acepta que estas profesiones están comprometidas con una misión particular, que incluye la necesidad de tomar decisiones profesionales que tienen que ver con cuestiones fundamentales de la existencia humana, como son: la vida, la propiedad y la libertad. Implica que los profesionales que representan estos campos disciplinares están en deuda o comprometidos con la confianza pública y deben actuar *pro bono*, por el bien público. Frente a un conflicto de intereses entre el interés público y el de una persona representando la profesión, el bien del cliente y del público se prioriza (sobre el bien personal).

4.6 SÍNTESIS CAPITULAR

A lo largo de este capítulo se enfatizó que el vínculo entre la evaluación y la ética es más fuerte de lo pensado. A este tenor, los principios y recomendaciones éticas propios de las organizaciones, siempre serán útiles para enfrentar los dilemas de la evaluación, de la toma de decisiones

colectiva, de las relaciones entre distintos actores involucrados y de los roles desempeñados. Indudablemente, el ambiente político es una condición que aporta complejidad e incertidumbre, pues los intereses que se representan imprimen presión a las actividades evaluativas y a la asignación de los recursos recibidos. La ética proporciona lineamientos muy importantes para que los equipos evaluadores enfrenten condiciones arrolladoras y garanticen la pulcritud exigida.

REFERENCIAS

- Boston, J.; Bradstock, A. y Eng, D. (2010). Ethics and public policy, en: Jonathan Boston, Andrew Bradstock y David Eng (eds.), *Public Policy: Why ethics matters*. ANU Press, 1-18. <http://www.jstor.org/stable/j.ctt24h2rv.5>
- Breczko, A. y Breczko, A. (2018). Ethics and the Challenges of the Contemporary World (Selected Problems), en: Miles Ventura (ed.), *Ethics*. Nova Science, 1-24.
- Cortina, A. (2005). El estatuto de la ética aplicada. *Hermenéutica crítica de las actividades humanas. Agora Philosophica*, VI(12), 7-22.
- Fitzpatrick, J. L.; Sanders, J. R. y Blaine R. W. (2004). *Program Evaluation: Alternative Approaches and Practical Guidelines*. Pearson Education Inc.
- Gardner, A. y Brindis, C. D. (2017). *Advocacy and policy change evaluation: Theory and practice*. Stanford Business Books.
- International Program for Development Evaluation Training (2007). *Evaluation Ethics, Politics, Standards, and Guiding Principles*. http://dmeformpeace.org/sites/default/files/M14_NA.pdf
- Khan, A. R. (2016). Politics, Ethics and Professionals. *Policy Evaluation. Journal of Social Sciences, Humanities, and Arts*, 16(2), 41-71.
- Mintrom, M. (2010). Doing ethical policy analysis, en: Jonathan Boston, Andrew Bradstock y David Eng (eds.), *Public Policy: Why ethics matters*. ANU Press, 37-54. <http://www.jstor.com/stable/j.ctt24h2rv.7>
- Neuman, W. L. (1997). *Social Research Methods. Qualitative and Quantitative Approaches*. Allyn and Bacon.
- Patton, M.Q. (1997). *Utilization-focused Evaluation* (3a. ed.). Sage.
- Rich, K. L. (2007). Introduction to Ethics, en: Janie B. Butts y Karen L. Rich, *Nursing Ethics Across the Curriculum and into Practice*. MA. Jones & Barlett Learning.
- Ryan, K. y Schwandt, T. A. (2002). *Exploring Evaluator Role and Identity*. Information Age.
- Scriver, M. (1969). An introduction to meta-evaluation. *Educational Products Report*, 2, 36-38.
- Scriver, M. (2009). Meta-evaluation Revisited. *Journal of MultiDisciplinary Evaluation*, 6(11), iii-viii.
- Sullivan, E. y Segers, M. (2007). Ethical Issues and Public Policy, en: Frank Fisher, Gerald J. Miller y Mara S. Sidney, *Handbook of Public Policy Analysis. Theory, Politics, and Methods*. CRC Press.
- Unsicker, J. (2012). *Confronting Power: The practice of policy advocacy*. Kumarian Press.
- Weiss, C. (1973). Where Politics and Evaluation Meet. *Evaluation Practice*, 14(1), 93-108.
- Weiss, C. (1998). *Evaluation: methods for studying programs and policies*. Prentice Hall.

CAPÍTULO 5

Evaluación basada en la teoría

INTRODUCCIÓN

Una vez revisadas las cuestiones particulares de la evaluación como un campo disciplinario, a partir de este capítulo se atenderán las distintas modalidades disponibles para la evaluación. Es importante insistir que, dependiendo del concepto de política asumido, la evaluación determinará su foco. Este capítulo se centra en las modalidades de la evaluación sobre los aspectos teóricos de una intervención (proyecto, programa o política). La evaluación teórica, en este sentido, parte de la idea de que una política pública o programa gubernamental es y está basado en una teoría o una lógica para atender un problema social o para alterar un determinado *statu quo*. De manera simple, puede afirmarse que una política pública es una teoría social del cambio. En consecuencia, una evaluación de esta naturaleza verificará la solidez de su propuesta (teoría, lógica o narrativa) para atender el problema en cuestión.

Entonces, la evaluación de esta naturaleza debe dar cuenta del soporte teórico. Esto es, de la congruencia de la teoría asumida y de su consistencia lógica para paliar o eliminar el problema, pues de otra manera no podría ser factible ni mucho menos tendría posibilidad alguna de éxito. Para cualquier evaluador profesional esto sería claro, pues no puede asumirse de entrada que:

1. La política tiene un soporte teórico.
2. Dicho soporte es el mejor o el más adecuado para resolver lo que intenta resolver. Por tanto, para esta modalidad evaluativa el foco es el diseño de la política y sus posibilidades científicas para enfrentar, de forma adecuada, la situación problemática. En consecuencia, una evaluación teórica verifica si existe correspondencia entre los propósitos perseguidos y los recursos dispuestos para tal fin. Asimismo, se pueden identificar varias opciones de la evaluación teórica.

De momento, estas modalidades se encuentran en desarrollo y no todas cuentan con un cuerpo teórico consolidado y aunque su importancia para algunos se considera marginal,

constituyen un referente importante para la verificación de la coherencia interna del programa o política pública.

5.1 LA SUPREMACÍA DEL DISEÑO EN EL PROCESO DE LAS POLÍTICAS

Por muchos años, los analistas creyeron que el éxito de una política se basaba en un buen diseño. Había que esmerarse en ello, era el supuesto, y lo demás fluía de manera correcta hasta lograr los objetivos deseados. Hoy por hoy, esta idea se ha desdibujado y los estudiosos del área encuentran determinantes adicionales e incluso más críticos para el buen desempeño y el éxito. Efectivamente, como lo afirma Howlett (2011, p. 30) “el diseño es crucial y buena parte de este es un determinante del éxito o fracaso de la política” pero, en definitiva, no lo es todo. A lo largo de décadas, el diseño ha tomado distintos rumbos y ha hecho diversos énfasis. Cada generación de estudios ha incrementado la complejidad del análisis, pero la gran observación es que el diseño de políticas y sus resultados están condicionados y determinados por los factores contextuales relacionados con la capacidad del gobierno, a la luz de los distintos niveles de complejidad social (Atkinson y Nigol, 1989, citado en Howlett, 2011). En épocas recientes, estas premisas se han movido de los asuntos contextuales, integrando las lógicas políticas e institucionales hacia las modalidades de *governance*. A pesar de la evidente importancia del diseño, la realidad social y política muestra que, en muchas ocasiones, se omite y que no todos los diseños son adecuados. Es claro que un diseño de política puede estar bien o pobremente hecho, ya que depende de muchas condiciones como las habilidades, el conocimiento y otros recursos dispuestos para la tarea de diseñar.

Desde la perspectiva del positivismo, la corriente tradicional y preponderante, el diseño de las políticas está basado en el concepto de cambio sistemático, de transformación de una situación indeseable, problemática o disfuncional, para convertirla en una situación deseada. En esta perspectiva subyace la imperfección de la realidad y la posición crítica del cambio y el mejoramiento. Es mucho el avance realizado y, en la actualidad, existen diversas técnicas que permiten el conocimiento y la reflexión para abordar la realidad y lograr dichas transformaciones. El diseño o la formulación tienen la misión de estructurarlas. En el estudio de las políticas, el diseño implica entender las herramientas disponibles en ese intento por alterar algunos aspectos de la sociedad y la conducta de los individuos. Estas herramientas son utilizadas para lograr algo o todo, de varias tareas intervinculadas en la formulación de las políticas (Jordan y Turnpenny, 2015). Tales tareas son:

1. Caracterizar el problema.
2. Evaluar el problema.

3. Especificar objetivos.
4. Identificar y evaluar opciones de política.
5. Diseñar la política. Para el diseño son muchas las consideraciones que los analistas asumen y los instrumentos que utilizan para apoyar sus propuestas, pero, para muchos, una buena teoría es lo más importante.

Desde el giro argumentativo, el diseño de una política es la resultante de la interacción humana, del uso del lenguaje y del ejercicio deliberativo efectuado en torno al diseño y, en general, al proceso de las políticas. Bajo esta perspectiva, el uso del lenguaje y la argumentación constituyen dimensiones esenciales de la teoría y del análisis (Fisher y Gottweis, 2012). Dado que esta perspectiva desafía la creencia de que el proceso de las políticas pudiera ser el producto de la aplicación directa del método y las técnicas científicas, propone también una perspectiva para: entender la relación entre lo empírico y lo normativo, reconstruir lo que el analista hace, y “entender cómo sus hallazgos son comunicados y cómo su consejo es entendido por quienes lo reciben” (Fisher y Gottweis, 2012, p. 3).

Las probabilidades de éxito en el tratamiento de la situación problemática que se enfrenta pueden (deben) crecer. El diseño, según la perspectiva tecnocrática, es un buen determinante del éxito. Independientemente de la perspectiva que lo oriente, lo cierto es que resulta el pilar de las acciones a emprender y constituye la base de las futuras consecuencias y logros obtenidos en la acción colectiva y de las decisiones de los responsables de los procesos políticos. Sin duda, se deben hacer los esfuerzos necesarios para que el diseño tenga la mejor estructura y se reduzcan las probabilidades de fracaso al mínimo posible.

5.2 LA POLÍTICA PÚBLICA COMO TEORÍA

Pressman y Wildavsky (1973) fueron los primeros autores en considerar una política pública como una teoría o una hipótesis causal, una cadena causal entre las condiciones iniciales y las futuras consecuencias: si x entonces y (Aguilar, 2000). Así las cosas, cualquier política o programa equivale a la formulación de una hipótesis sobre las acciones humanas o gubernamentales que causarían cadenas de futuros eventos, transformando una situación problemática. Por ejemplo, si un gobierno otorga un subsidio a las familias y si estas usan adecuadamente los recursos, sus miembros mejorarían su calidad de vida e ingresarían a un mercado en condiciones favorables. Esta perspectiva, sin duda, refleja la visión positivista de una política pública, al asumir una visión científica del asunto a tratar: *si... entonces...* La causalidad asumida por el método científico se revela en toda la extensión de la palabra y bastaría con identificar las mejores causalidades o cadenas causales que condujeran a un objetivo deseado.

Bajo esta perspectiva, se observa también la situación tentativa de la formulación de la política; por tanto, una política será una hipótesis por verificar en la realidad y, en conse-

cuencia, encarnaría de entrada un margen de error, dada la condicionalidad inherente. Esta faceta tiene implicaciones importantes para la acción pública y para el acto de gobernar. De otra forma, la solución de problemas sociales nunca podría asegurarse con anterioridad, mucho menos cuando sus postulados se basan en preferencias ideológicas, visiones personales o tradicionales. Si la visión científica tiene algún margen de error, este puede ampliarse aún más cuando sus supuestos son débiles, incoherentes o falsos. Entonces, todo conocimiento sería parcial, limitado y de ninguna manera absoluto o general. Dos sentencias derivan del conocimiento: la ciencia tiene sus límites y los problemas de vivir en sociedad adquieren, en ocasiones, la dimensión de problemas perniciosos (*wicked problems*).

Para los propósitos de la evaluación, verificar las bases teóricas de la política pública resulta ser imperativo. Se supone entonces que la consistencia de la teoría y su solidez reducirían al mínimo el margen de error en la atención del problema identificado y que se intenta superar. El supuesto es simple: si la teoría para la solución de mi problema es acertada, el problema se resolverá. Es así como la evaluación encuentra su foco inicial en la teoría, la hipótesis o la causalidad que sustenta o soporta la política; en concreto, en la coherencia de sus supuestos, valores y estrategias y en su validez científica. Por tanto, el desafío inicial del evaluador es verificar si la teoría elegida es la adecuada. Subyace la duda de si los *policymakers*, decisores y todos los responsables toman las mejores decisiones.

5.3 UN BREVE PARÉNTESIS ACERCA DE LA TEORÍA Y LA CAUSALIDAD

Una teoría resalta y explica algo que uno de otra manera no observaría o simplemente encontraría enigmático, intrigante o misterioso (Gilbert, 2001). A menudo, la teoría es una respuesta a una pregunta: ¿por qué hay gente pobre y otra rica?, o ¿por qué hay desempleo? La teoría social se define como un sistema interconectado de abstracciones o ideas que condensan y organizan el conocimiento acerca del mundo social (Neuman, 1997). Es esta teoría una manera compacta de pensar el mundo social y su característica principal es la posibilidad de utilizarse como explicación de un fenómeno o suceso. Por lo regular, a las teorías se llega por inducción, es decir, extractando los elementos comunes de instancias específicas a la realidad y aplicándolas después, para explicar o comprender otras instancias, utilizando la lógica deductiva.

En sus cometidos, teorizar exige la conceptualización y validación de las hipótesis de trabajo como también de las declaraciones o posiciones del investigador. A la par, para probar una teoría es recomendable seleccionar indicadores para los respectivos conceptos que permiten hacer comparaciones entre la información recolectada y las predicciones extraídas. En la investigación social es conveniente decirlo, ya que se encuentran teorías con distintas posibilidades para explicar los fenómenos sociales bajo estudio. Significa esto que existen unas

teorías más adecuadas que otras, unas con más poder explicativo que otras. Es usual reconocer conceptos, supuestos, clasificaciones, relaciones y espectros como componentes de una teoría.

Un punto de partida para generar una teoría es adoptar una perspectiva epistemológica, elección esta que condiciona la manera de concebir el mundo y los fenómenos. La explicación causal es la manera más simple y común de explicar un fenómeno, de modo que el positivismo la asume como su principal estrategia. Dado que el positivismo no considera el azar como razón válida, los eventos suceden porque algo los causa. Desde esta perspectiva nada es aleatorio y al investigador le corresponde identificar los hechos generadores o precedentes que a la postre explican su ocurrencia. En el positivismo, la posición ontológica asume que la realidad es externa, regular y que, sencillamente, debe ser descubierta. De manera simple, la causalidad expresa una relación entre dos hechos: uno precedente y otro consecuente.

Según los autores, tres son las condiciones asumidas como básicas de la explicación causal:

1. El orden temporal establece que la causa debe venir antes del efecto, de modo que se establece la dirección de la causa hacia el efecto.
2. La asociación que debe reconocer el investigador, “dos fenómenos están asociados si ocurren bajo un mismo patrón o si parecen actuar juntos” (Neuman, 1994, p. 51). Dicha asociación es fuente de confusión, ya que algunos asumen una correlación como asociación y esta con causalidad.
3. La eliminación de alternativas significa que el investigador debe demostrar que el efecto es debido a la variable causal y no por otra(s) circunstancia(s). Desde esta óptica se evitan relaciones espurias. La investigación experimental es el medio idóneo para estudiar la eliminación de alternativas (inobservable e indirecta).

Al positivismo se le critica, de manera usual, por la manera como concibe la ocurrencia de los hechos, asumiendo para ello la linealidad o transitividad con que estructura una explicación causal: A causa B, B causa C, C causa D; constituyendo esto una aparente debilidad por la forma de concebir los fenómenos. Aun así, suele emplearse para abordar fenómenos sociales utilizando proposiciones simples y explicaciones causales lineales con una o varias variables. Por ejemplo, el desempleo podría explicar la inseguridad o las diferencias salariales debido al género del trabajador, usando al inicio la deducción y, después, probando relaciones empíricas (datos en la realidad). Podría plantearse también que x causa y , o que y ocurre por x , donde x y *problemas* son conceptos o variables. Algunos investigadores pueden plantear relaciones de manera predictiva: si x , entonces y o dado un x , entonces se produce un y . Otros, con mayor cautela, proponen que dado x , entonces probablemente ocurrirá y . Es conveniente precisar que las relaciones entre las variables pueden ser positivas o negativas. Positiva, es decir, un aumento en una variable conlleva aumento en la misma dirección de la otra varia-

ble; a mayor salario se esperaría un mayor nivel de ahorro. Negativa (inversa), un aumento de la primera variable representa una disminución de la segunda variable; a mayor nivel educativo se esperaría una mejor planeación del número de hijos. Estas son suposiciones teóricas o hipótesis de trabajo a validarse; la realidad puede demostrar lo contrario.

5.4 LA EVALUACIÓN BASADA EN LA TEORÍA

La literatura especializada presenta distintas modalidades de evaluación dedicada a los asuntos de la teoría, todas ellas agrupadas en los que regularmente se reconoce como evaluación basada en la teoría (*Theory-based Evaluation* [TBE, por sus siglas en inglés]). Este conjunto de posibilidades es una colección de métodos que enfatizan en la importancia de los supuestos de cómo y por qué un programa (política) funciona, a fin de evaluarlo (Breuer *et al.*, 2016). Estas incluyen: la Teoría del Programa (*Program Theory*); el modelo lógico (*Logic model*); la Teoría del Cambio (*Theory of Change*); la Cadena de Resultados (*Result chain*); la Ruta de los Resultados (*Outcome pathway*); la Teoría de la Acción y la Teoría de la Implementación (*Implementation Theory*). Estas modalidades de evaluación tuvieron su origen en las décadas de los años sesenta y setenta, cuando los científicos sociales y los estudiosos de las políticas públicas propusieron modelos para representar las relaciones entre las variables de los fenómenos de estudio (Leeuw, 2012). Fue una manera de otorgarle carácter científico a las propuestas y de imprimirle racionalidad a las decisiones. No comprende un campo compacto ni puede hablarse de una única opción y se asimila a una familia de estrategias utilizadas para ilustrar cómo las intervenciones gubernamentales son pensadas.

Estas posibilidades usualmente hacen referencia a “modelos” que guían la acción de las organizaciones o agencias para transformar la realidad. Todas las modalidades de evaluación aquí presentadas comparten los mismos afanes y persiguen los mismos propósitos: establecer el grado de coherencia interna, estudiar cadenas causales y dar seguimiento a la concepción del cambio, todo bajo una perspectiva comparada entre lo teórico y lo empírico. Estas posibilidades intentan corregir lo procedimental de la evaluación, ya que en ocasiones los evaluadores ponen más atención al proceso de evaluación, que al mismo programa o a sus fundamentos y principios. Ellos dan por hecho que el programa o la política están bien diseñados, lo que resulta a todas luces inconveniente. En consecuencia, una de las recomendaciones para abordar un proceso de diseño de evaluación es prestar mayor atención a la estructura del programa, estudiando sus premisas teóricas y verificando su racionalidad. El equipo evaluador es imperativo, pues debe familiarizarse con el campo y aprender acerca de lo que en particular está pasando con el programa bajo estudio. El analista debe entender con exactitud cómo se supone que el programa inducirá el cambio deseado. La evaluación dedicada a los supuestos teóricos se reconoce, entonces, como un ejercicio valioso y fructífero.

5.4.1 Evaluación lógica

La evaluación lógica es entendida como un “modelo plausible y sensible sobre el supuesto funcionamiento de un programa” (Bickman, 1987 en Wholey, 2010, p. 56). Es el conjunto de supuestos entre las estrategias y tácticas adoptadas por el programa y los beneficios sociales que se esperan alcanzar (Rossi *et al.*, 1998). Se define como “los mecanismos que median entre la provisión y recepción del programa y la emergencia de resultados de interés” (Weiss, 1998, p. 57). En la base de estas definiciones se encuentra que los programas o las políticas descansan en un esquema donde la causalidad determina su diseño: si... entonces... Desde lo operativo, la evaluación lógica intenta conformar la cadena de supuestos causales para vincular recursos, actividades, efectos intermedios y metas del programa (Wholey, 1987 en Wholey, 2004).

Por ello, conviene definir desde el comienzo las metas oficiales del programa; esto es, conocer la intención del programa. En efecto, la mayoría de los programas tienen una declaración escrita que describe actividades y metas de forma clara, pero otros las presentan de manera confusa y ambigua. Cuando el programa busca alcanzar metas oficiales, pero a su vez asuntos no declarados (soterrados), con toda seguridad se enfrentará una dificultad. De igual modo, es de reconocer que los programas pueden tener múltiples metas a dilucidar y jerarquizar. Es probable que observando el programa en operación o hablando con el personal de servicios y con los beneficiarios se entiendan los verdaderos propósitos.

Para evaluar resulta útil saber no solo lo que el programa aspira alcanzar, sino también los medios estipulados para lograrlo. Por tanto, lo óptimo es entender las premisas teóricas que soportan el programa, ejercicio fructífero si el evaluador investiga y entiende la teoría implícita o explícita del programa. La teoría no tiene que ser aceptada de manera uniforme o ser correcta, en principio serán hipótesis que soportan un plan de acción.

Las reflexiones acerca de las relaciones entre variables posibilitan la estructuración de diagramas causales como medio para abordar y entender problemas más complejos y para emprender una evaluación teórica o lógica. Al más simple nivel se tienen los diagramas causales entre dos variables, y para más de dos variables y relaciones de distinto orden se disponen los diagramas multicausales. Estas herramientas permiten abordar fenómenos complejos y estructurar teorías abstractas de forma compacta. Bajo esta perspectiva, muchos *policymakers* proponen discursos tecnocráticos para atender las disfuncionalidades que a menudo se presentan en las sociedades. En consecuencia, el evaluador en primera instancia debe revisar la estructura causal subyacente.

5.4.2 Evaluación teórica

La otra denominación de esta posibilidad evaluativa es reconocida como evaluación teórica. Para Wholey (1987, citado en Wholey, 2010: 56), la teoría del programa identifica recursos, actividades y resultados intencionados y especifica una cadena causal de supuestos vincula-

dos, recursos, actividades, resultados intermedios y metas últimas. La teoría del programa como la utiliza Weiss (1998), se refiere a los mecanismos que median entre la provisión y recepción del programa y la emergencia de los resultados de interés. Los mecanismos operativos de cambio no son las actividades del programa *per se*, sino la respuesta que las actividades generan. Estas o cualquiera de las muchas respuestas (cognitivas, afectivas, sociales o culturales) serían los mecanismos que conducen a los resultados deseados. *Program Theory* pone el énfasis sobre las respuestas de la gente a las actividades del programa. Usualmente incluye: *inputs* y auspicios organizacionales, actividades del programa, resultados intermedios, y resultados finales. Este ejercicio es útil no solo para el evaluador, sino también puede ayudar a diseñadores de programas, hacedores de políticas públicas, practicantes y patrocinadores.

Las ventajas derivadas de tal ejercicio son muy convenientes, pues permiten la generación de un buen sentido de los asuntos inmersos en la iniciativa, generando mayor criterio y también controversias derivadas. En otro ámbito, una vez establecida la teoría del programa, permite formular las preguntas orientadoras del posterior proceso de evaluación y, a la par, alienta el entendimiento de los datos y proporciona elementos para entender la evidencia. Si se conoce la lógica del programa, esto, de alguna manera, otorga autoridad para hacer recomendaciones sobre la vida futura del emprendimiento y en lo funcional, y determina los elementos relevantes para el reporte y sustentación de la función de evaluación. Hay una serie de pasos para familiarizarse con el área del programa y con sus especificidades que van desde la lectura de temas relacionados con el campo de intervención, como el sondeo en los antecedentes y evaluaciones previas, o bien, con el contacto y diálogo con la gente del programa.

La evaluación teórica tiene distintas posibilidades para llevarse a cabo, pero el lenguaje matemático y los algoritmos computacionales se han convertido en importantes medios para la evaluación teórica en la ciencia política (Grant, 1973). Según este autor, el proceso implica:

1. Trasladar las principales relaciones teóricas a un lenguaje formal para el análisis de su consistencia interna.
2. Interpretar empíricamente el lenguaje formal a fin de hacer predicciones específicas, las cuales, a su vez, permiten la evaluación de la consistencia externa con el significado teórico del fenómeno del mundo real.

Más allá de estas consideraciones iniciales, conviene justificar aquellos asuntos de la política como teoría del cambio social, sus primeros esbozos y las consecuencias derivadas.

5.4.3 Teoría del cambio

En 1995, Carol Weiss popularizó el término “teoría del cambio” y argumentaba que la mayoría de las veces las políticas eran difíciles de evaluar, porque estaban pobremente articuladas. Su reclamo fue más que justificado, a la luz de las experiencias acumuladas en el ejerci-

cio como evaluador o analista de políticas públicas. En realidad, la ausencia de claridad y de la secuencia detallada de cómo se supone el cambio tiene su origen en una de las características de las políticas y programas puestas en escrutinio. Por lo regular, los servidores y responsables del diseño asumen, de manera ingenua, que el cambio ocurre “por decreto” o por la popularidad de los gobernantes. Pocas veces consideran el largo plazo como un horizonte de tiempo. En virtud de ello, la propuesta de la famosa evaluadora (teoría del cambio) buscaba subsanar mucho de estas imprecisiones al constituir una vía para describir el conjunto de supuestos que explican los minipasos para alcanzar resultados a largo plazo. Pero también las conexiones entre las actividades de la política y los resultados obtenidos en cada paso.

Una revisión reciente refiere la teoría del cambio como una herramienta o perspectiva para reflexionar sobre cómo y por qué podría ocurrir el cambio (Vogel, 2012). Exige un pensamiento lógico y una reflexión crítica sobre los contextos en que ocurren ciertos hechos a modificar. Para la mayoría de los exponentes, la participación es una condición, ya que la teoría del cambio es un proceso tanto reflexivo como dialógico entre similares, con *stakeholders*, en el cual se reflejan los valores, visiones y filosofías del cambio. Los individuos, en dicho proceso, explicitan sus creencias y supuestos sobre cómo y por qué el cambio podría darse como resultado de una iniciativa: si hago esto... entonces... Por ello, para algunos, la teoría del cambio mapea la secuencia lógica de una iniciativa desde sus insumos hasta sus resultados. La combinación de la lógica con el pensamiento crítico, en un contexto determinado, permite conocer las motivaciones, contribuciones e interpretaciones (supuestos) de los actores acerca de cómo podría aparecer el cambio.

De manera particular, la teoría del cambio se traduce en “la descripción de una secuencia de eventos esperada para conducir a un particular resultado deseado” (Rick Davis, 2012, citado en Vogel, 2012, p. 9). Lo cierto es que cualquier ejercicio bajo esta modalidad amplía la comprensión del cambio, reconociendo su complejidad y permitiendo elaborar una narrativa y una representación diagramática de cómo ocurre. Para las Naciones Unidas, una teoría del cambio es un método que explica cómo se espera que una intervención o conjunto de intervenciones conduzcan a un cambio específico en el desarrollo, la cual está fundada en el análisis causal y basada en la evidencia disponible.

Quienes buscaban una reflexión consciente sobre las teorías del desarrollo, establecen los orígenes de la teoría del cambio en los enfoques participativos de los años sesenta del siglo XX. Se reconoce como un aspecto del *program theory* desde donde se exigían los fundamentos teóricos de un programa y se urgía por una clara articulación entre insumos y resultado. Bajo estas premisas y exigencias, la teoría del cambio hace parte de las teorías del cambio social desde donde se ha debatido sobre aquellos asuntos que conducen al desarrollo. Otros mencionan al educador brasileño Paulo Freire, quien en 1970 ayudó a las personas a concientizar sus creencias acerca de la pobreza y cómo tratarla y a reflexionar sobre ello y tomar acciones. Lo veía como un proceso de empoderamiento en sí mismo. En 1990, un grupo de escritores desafiaron a los programas al articular los cambios que querían lograr y lo

que debía pasar para consolidarlos. La mesa redonda sobre el cambio de la comunidad del instituto de investigación Aspen se unió con ActKnowledge, una organización independiente para la construcción de capacidad y producir así la primera guía sobre la teoría del cambio (James, 2011).

La teoría del cambio, entonces, puede ser aplicada en los diferentes niveles, etapas o fases del ciclo de la política. Ella sirve para clarificar los determinantes, tanto internos como externos, del éxito o del fracaso. Asimismo, identifica los distintos *stakeholders*: donantes, operadores, ejecutores, fundaciones, sociedad civil, organizaciones para el desarrollo y todos aquellos involucrados en torno a la iniciativa. Por supuesto, también sirve como marco de referencia desde el cual se evalúa el impacto y se mejora el monitoreo. Sobre todo, sirve para probar los supuestos, para demostrar el impacto y para generar teoría. En general, aquellos que la impulsan, la recomiendan para fortalecer la gerencia adaptativa. Evaluar una teoría del cambio implica pensar en la calidad del proceso de reflexión, los conceptos subyacentes, la claridad del proceso de pensar y su utilidad. El proceso de reflexión, en todo caso, debe ser participativo al extremo, inclusivo, situado y profundo. El producto de una intervención basada en esta teoría suele ser un documento que da cuenta de la discusión, la narrativa y el mapa o estructura lógica. En este documento deben exponerse los “mecanismos causales” considerados en una intervención y que proponen la combinación de procesos sociales y contextuales a activar para lograr los cambios deseados. La “inferencia causal” requiere de evidencia e información para demostrar la relación entre los cambios ocurridos y las actividades que el programa emprendió. Más allá de la demostración de los resultados e impactos, conviene indagar también acerca del aprendizaje logrado, a fin de mejorar las estrategias e intervenciones, como también sobre la rendición de cuentas para donantes y para la comunidad local.

La discusión fomentada por la teoría del cambio debe considerar como mínimo, asuntos sobre el contexto de la iniciativa, sus ámbitos de acción, el cambio a largo plazo y la secuencia para que ocurra el cambio. Finalmente, reflexionar sobre los supuestos para que suceda el cambio y verificar si las actividades y los resultados orientan hacia el cambio.

5.4.4 Evaluación basada en la teoría (EBT)

Campo disciplinar con desarrollos importantes en los últimos 40 años y alternativa para analizar complementariamente ciertas intervenciones. Figuran como características de la EBT (*Evaluation Based Theory*):

- a) Enfoque en los supuestos e hipótesis, asumidos por el programa, si... entonces...
- b) dos momentos, uno conceptual que explica la teoría del programa o modelo de la política, y otro empírico que indaga la forma en que el programa opera y logra los resultados,

- c) identificación de los mecanismos, medios, estrategias, incluso actividades y vínculos entre el contexto y los resultados (Pawson, 2006; Pawson y Tilley, 1997, en Leeuw, 2012). Los mecanismos son “una constelación de entidades y actividades vinculadas entre sí, de manera que conducen a un particular tipo de resultado” (Hedström, 2005 en Leeuw, 2012, p. 349). En términos simples son aquellos recursos, actividades, estrategias, iniciativas y medios necesarios para alcanzar un propósito, y
- d) aplicable a diferentes métodos para reconstruir y probar supuestos, hipótesis o lógica que subyacen al cambio.

5.4.5 Análisis de la contribución

Desarrollada por John Mayne a principios del año 2000, el Análisis de la Contribución (*Contribution Analysis*) es una metodología que identifica los aportes de una iniciativa de desarrollo (proyecto, programa, política) para lograr el cambio o un conjunto de cambios en una sociedad o grupo social. Subyace el reconocimiento de las dificultades para estudiar la causalidad, identificar sus efectos y atribuir tales cambios a las intervenciones adelantadas por una autoridad, una organización o un grupo de *stakeholders*. Es un simple reconocimiento de la complejidad de la acción colectiva, de los múltiples factores que pudieran afectar los resultados de cualquier iniciativa y que difícilmente establecen su verdadera conexión con el programa emprendido. Eso es debido a que el cambio requiere muchos pasos, así como actividades intermedias diferentes para alcanzar el cambio deseado, pues no es directo ni automático (Mayne, 2012). También debe reconocerse que, en cualquier emprendimiento de desarrollo social, influyen factores externos y muchas intervenciones simultáneas pueden contribuir a un simple cambio.

El análisis de la contribución, a la par de las anteriores propuestas que evalúan la teoría, está interesado en establecer las conexiones atribuibles entre la iniciativa llevada a cabo y el cambio alcanzado. Aspira a establecer cómo el cambio ocurre o se supone que ocurre en sus distintos niveles o cómo las actividades conllevan a ciertos resultados intermedios, a resultados finales y provoca impactos finales; esto se realiza para comparar la teoría con la realidad. La meta es producir una narrativa plausible, basada en evidencia con la cual una persona normal pudiera concordar para explicar cómo y por qué ocurrió el cambio (Mayne, 2012). Esta narrativa se considera válida si se basa en una teoría sólida del cambio, con supuestos que explican de manera razonable cómo el programa o la política propician el cambio deseado. Obviamente, esta narrativa da cuenta que las actividades fueron implementadas como estaban planeadas y llevadas a cabo de manera apropiada. Esta narrativa es válida, ya que hay evidencia del cambio en cada nivel de la teoría del cambio, y puede constatarse que la relativa influencia de los factores externos o de otras intervenciones es demostrable.

Establecidas las principales pretensiones del análisis de la contribución, el hecho de establecer la causalidad de cualquier evento implica un desafío mayúsculo en cualquier ámbito.

Tales retos son todavía más exigentes en las intervenciones del ámbito social que aspiran a mejorar la condición de sus agentes, donde la complejidad prevalece. Aún más difícil es establecer aquella contribución que convoca muchas iniciativas, distintos actores con intereses contrapuestos y con eventos inesperados a la orden del día. Sin embargo, cualquier esfuerzo hecho es enteramente válido y necesario, pues la validación de la teoría es una responsabilidad permanente del ejercicio investigativo.

5.5 INSTRUMENTOS QUE SOPORTAN LA EVALUACIÓN TEÓRICA/LÓGICA

Cada una de las perspectivas evaluativas centradas en los soportes teóricos de las políticas presenta sus propios desarrollos metodológicos para facilitar su implementación y guiar a los equipos evaluadores. A continuación, se presentan las particularidades del modelo lógico de la Fundación W.K. Kellog, lo propio de la propuesta de la evaluación teórica de Carol Weiss y el árbol de problemas desde la experiencia de Chevalier. La intención es develar sus particularidades, a fin de facilitar su uso y mejoramiento.

5.5.1 El modelo lógico

La Fundación W.K. Kellog (2004), mediante su propuesta, vincula los supuestos teóricos subyacentes del programa con los componentes a desarrollar durante su implementación. “El modelo lógico presenta visual y sistemáticamente las relaciones entre los recursos disponibles para operar el programa, las actividades previstas y los cambios o resultados que se esperan alcanzar” (W.K. Kellog Foundation, 2004, p. 2).

Inicialmente, son dos los componentes del modelo: el trabajo planeado y los resultados perseguidos (figura 5.1). Estudiar y evaluar la lógica de una política requiere, según esta visión, identificar la secuencia entre lo que se planifica y lo que se aspira a conseguir, los recursos necesarios y las actividades requeridas para obtener resultados tangibles (*outputs*), impactos (*outcomes*), así como algunas alteraciones notables en el bienestar de las personas involucradas. En este planteamiento se aprecia una conexión explícita entre las etapas de planeación y evaluación. En ocasiones, los hacedores de política actúan en función de diferentes motivaciones ideológicas, dejando de lado, la reflexión rigurosa de los supuestos y los fundamentos teóricos y lógicos que soportan la propuesta de intervención.

La figura 5.1 orienta al evaluador procedimentalmente sobre asuntos a indagar, énfasis y focos de labor y descubrimiento de premisas y principios de la política. Por esta razón, las iniciativas públicas no pueden soportarse en promesas ligeras de campaña, sin reflexión y animadas por el deseo de triunfar en la política. Cuando el equipo somete a evaluación el programa o la política pública, se revelan las verdaderas intenciones del político en el poder. En ocasiones, la dinámica del sector público obliga la ejecución de programas con recursos

insuficientes, y es que la premura de la acción lleva a iniciarlos y buscar, en paralelo, los recursos necesarios en la medida en que las circunstancias sean favorables. Asimismo, procede enlistar todas las actividades requeridas para lograr el cambio y obtener resultados intermedios, cambios e impactos. Es evidente que la causalidad se hace presente en este modelo de planeación y evaluación, quedando claro que, para lograr el cambio, inducir o generar nuevas condiciones y relaciones, se deben emprender procesos sistemáticos, rigurosos y participativos.

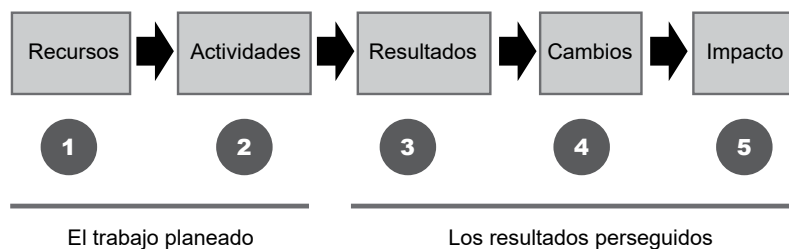


FIGURA 5.1. MODELO LÓGICO BÁSICO

Fuente: W.K. Kellogg Foundation, 2004, p. 1.

5.5.2 La evaluación teórica de Weiss

Carol Weiss propuso un enfoque para dar tratamiento a las cadenas causales que pudieran reflejar los supuestos teóricos y secuenciales para lograr las transformaciones deseadas. En el marco de esta modalidad, la teoría debe tomarse en serio, desarrollarse de manera rigurosa e, incluso, considerar el diseño experimental. Solo así, dice la autora, se contribuirá a la investigación científica genuina. La observación sistemática y el monitoreo permanente ayudan a comprender los alcances y población objetivo del programa y, sobre todo, a determinar el apego a la planeación de las actividades y la correcta utilización de los recursos.

Weiss encuentra que la teoría del cambio ha sido crucial para el desarrollo de la evaluación de los programas; por ello, recomendaba a los evaluadores entender el programa y determinar precisamente cómo funcionaba. “Muchos programas son difíciles de evaluar, en esencia, porque teóricamente son pobres y sus supuestos presentan una articulación débil” (Msila y Setlhako, 2013, p. 324). El supuesto es muy simple, pero contundente para emprender un ejercicio evaluativo, si se conocen las teorías subyacentes a los programas ayuda en la evaluación, en la mejora y en la determinación de la teoría que soporta los resultados. El cambio pretendido debe tener un origen y no puede esperarse que sea resultado del azar.

Para inducir el cambio, se sugiere que cada comunidad tenga un mapa que ilustre los destinos y las rutas a seguir para lograr el progreso (*Organizational Research Services*, 2004, citado en Msila y Setlhako, 2013). Este mapa revela los supuestos, tales como el destino final,

el contexto imperante, los compromisos durante la jornada y el sistema de creencias que sustenta el emprendimiento. Sin una teoría de cambio resulta imposible una transformación deseada.

La figura 5.2 presenta un mapa causal, centrado en el sector educativo que pretende propiciar un mejoramiento en el desempeño de los estudiantes. El punto de partida elegido es el incremento salarial de los docentes y a partir de ello se establece la cadena causal. Se supone que una mejora en los salarios de los maestros elevaría su moral, los estimularía y, a su vez, los motivaría a trabajar más y dedicar más tiempo a su actividad. A la par, se esperaría una mejora de las relaciones con la comunidad estudiantil y con ello se despertaría un espíritu colaborativo que a la postre propiciaría una mayor dedicación y rendimiento de cada uno.

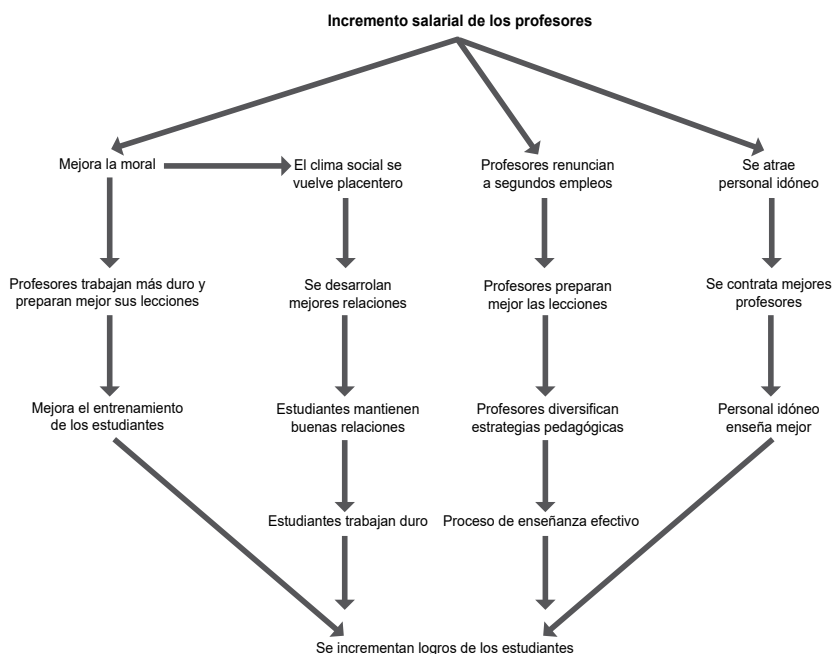


FIGURA 5.2. CADENA CAUSAL EN EL SECTOR EDUCATIVO

Fuente: Weiss, 1998, p. 63.

En otra cadena causal, una mejora del clima laboral y organizacional supone mejores relaciones entre los actores involucrados y un estímulo para el trabajo arduo de los estudiantes. Por otra parte, se supondría que los docentes dedican mayor tiempo y esfuerzo a una sola institución, fortaleciendo su identidad con el plantel y, consecuentemente, su trabajo mejoraría al diversificar estrategias pedagógicas, garantizando la efectividad del proceso enseñanza-aprendizaje. Un efecto adicional se daría en función de la calidad del sistema educativo, ya

que al ser atractivos los salarios, habría más personal capacitado. Como efecto final, serán evidentes el desempeño y los logros de las comunidades de estudiantes y docentes. Los salarios justos no pueden considerarse la panacea para mejorar la actividad docente y estudiantil, por lo que deben observarse otros factores determinantes del buen desempeño.

El anterior esquema puede considerarse teoría del cambio diseñada para conseguir un determinado objetivo. Al serlo, presenta en detalle los supuestos adoptados y la ruta del cambio con sus resultados intermedios hasta llegar a lo aspirado. Por supuesto, los efectos de una decisión están soportados en postulados provenientes de la psicología, la economía, la pedagogía y sobre el desarrollo social. Aquellos preocupados por escalar los logros estudiantiles y del sistema educativo encontrarán, en este esquema, un mapa detallado de cómo se supone el cambio que toma lugar para la consecución de los objetivos. La cadena causal presume solidez teórica, pero el evaluador deberá constatar la magnitud obtenida en los hechos. Significa también que diversas condiciones contextuales y/o personales pueden impedir el funcionamiento del modelo, por ejemplo, docentes que no reaccionen al incremento salarial o estudiantes que no valoren los esfuerzos adicionales. Asimismo, el modelo tendría un margen de error, la cadena causal; por tanto, este no es infalible y corresponde al evaluador establecer qué se puede atribuir a la teoría y qué otras razones ajenas a estos supuestos contribuyeron a ellos.

5.5.3 El árbol de problemas

Para formular un problema, se recomienda el uso de instrumentos que reduzcan su complejidad y diluciden sus estructuras; de ahí que el árbol de problemas se reconozca como estrategia analítica, basada también en la cadena causal y que facilita la comprensión de las causas y los efectos de un problema. Asimismo, permite identificar algunas alternativas de solución para el problema de una política.

El árbol de problemas es una heurística usada para la estructuración de problemas, cuando el analista enfrenta situaciones complejas y vagamente definidas (Vesely, 2008). También ayuda a analizar las causas y efectos de primer y segundo nivel de un problema central, permite comprender lo que causa un problema como: anticipar cuáles podrían ser sus efectos (Chevalier y Buckles, 2009). Su nombre deriva de un símil, ya que el problema central es el tronco vinculado a las ramas o efectos y tiene sus orígenes en las raíces (figura 5.3).

Para la estructuración de un árbol siempre es recomendable la participación permanente de los *stakeholders*, ya que permite contrastar distintas visiones de los involucrados. Los expertos recomiendan ilustrar los conceptos relacionados: problema, efecto y causa. Aunque no hay consenso sobre el punto de partida, la experiencia personal con la herramienta indica que es la expresión negativa observable o evidente (sentimientos, emociones y sensaciones), la que revela la presencia de un problema central: incomodidad, frustración, enojo, angustia, preocupación, etcétera.

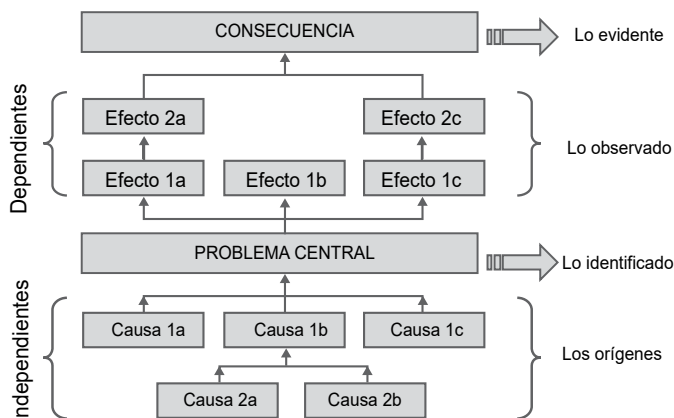


FIGURA 5.3. ESTRUCTURA DEL ÁRBOL DEL PROBLEMA

Fuente: elaborado con datos de Chavalier y Buckles, 2009, p. 129.

La observación de hechos recurrentes (efectos) son indicativos de que algo disfuncional ocurre, por ejemplo, si los ciudadanos llegan tarde a sus destinos, es indicativo de que “algo ocurre” durante el desplazamiento (figura 5.4). Estas situaciones pueden ser cuantificables y traducidas en tasas o indicadores: minutos/semana de retardos, heridos diarios, accidentes por minuto. El analista debe establecer razones subyacentes a estos eventos, emprender tareas de reflexión, indagación y validación tendientes a la definición del problema central. De nueva cuenta, la causalidad es la condición *sine qua non* para conectar el problema con los efectos; las situaciones ocurren dada una condición anterior.

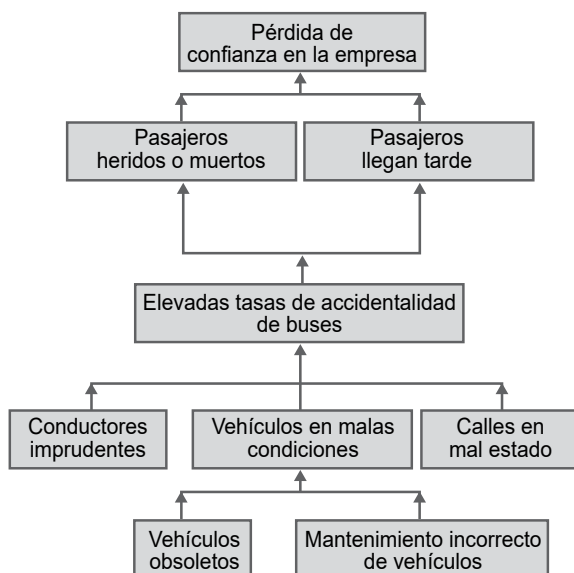


FIGURA 5.4. EFECTOS DE LOS ACCIDENTES DE TRÁNSITO

Fuente: Chavalier, 2004, p. 8.

Un problema, por tanto, es una situación disfuncional, incómoda e indeseable, que invita a superarla dadas las consecuencias generadas; por tanto, el analista debe preguntarse e indagar qué lo ocasionó. El reconocimiento de las causas no es un ejercicio especulativo y debe(n) presentarse la(s) evidencia(s) que soporta(n) los eventos. También es importante reconocer la existencia de distintos niveles de causas, estableciendo la jerarquización del caso y dando sentido a la cadena causal. En la figura 5.4 se aprecian dos niveles de causas, de modo que a una causa de primer nivel (vehículos en malas condiciones) se vinculan dos causas de segundo nivel (vehículos obsoletos y mantenimiento incorrecto de los vehículos).

5.6 IMPORTANCIA DE EVALUAR LOS SUPUESTOS DEL DISEÑO

El propósito ulterior de la evaluación teórica es argumentar la consistencia interna del programa y las posibilidades de generalización en la realidad. Para confirmar la consistencia interna (validez interna), uno de los aspectos relacionados es su representación, misma que no debe contener contradicciones internas. Esto es debido a que la principal razón del uso de representaciones matemáticas o logarítmicas es la deducción de nuevas proposiciones para evaluaciones empíricas. Las propiedades lógicas de la representación deben ser implícitas, de manera que los resultados de las pruebas de sensibilidad no se confundan con fuentes inherentes o espurias de la varianza. El entendimiento de los basamentos teóricos de un programa permite organizar la evaluación, rastrear los supuestos no evidentes y determinar los momentos apropiados para recolectar información. Por supuesto, mapear los supuestos del programa ayuda a descubrir los momentos en que se rompe la cadena causal, juzgando si se va por buen camino o por una ruta equivocada y explicando los pasos y procesos que llevan a los resultados esperados o hacia consecuencias inesperadas.

En ocasiones se sugiere el apoyo computacional, a fin de rastrear la teoría del programa y sus principales componentes. El test de sensibilidad es una herramienta básica para la evaluación de las teorías, que usa representaciones matemáticas o algorítmicas y permite la manipulación de algún componente de la teoría y observación de las consecuencias lógicas o sustantivas (Grant, 1973). En general, se duda que la gente sea tan buena identificando el proceso causal que está funcionando, que el evaluador tenga alguna idea de cómo ocurre el cambio o que el conocimiento científico social sea adecuado para identificar las formas como se da el cambio; una cosa es la teoría del programa y otra la implementación. El supuesto es que si las actividades son conducidas como fueron planeadas y realizadas con suficiente calidad, intensidad y fidelidad al plan, los resultados deseados serán alcanzados.

Estas modalidades evaluativas intentan determinar la evolución del programa y cada una de las fases de sus actividades, mismas que convocan una respuesta apropiada de los participantes, la cual condiciona las siguientes fases. A menos que la agencia realice correctamente las actividades programadas, no es probable que los participantes hagan progresos hacia la

meta. Si los participantes están desinteresados, sin voluntad o sin capacidad para aprender, o resistentes a tomar su trabajo, el progreso no tomará lugar. Entonces, el conocimiento del programa y toda su lógica subyacente permite identificar las preguntas claves para el estudio.

5.7 SÍNTESIS CAPITULAR

En el análisis de políticas públicas se suele(n) asumir alguna(s) teoría(s) que pueda(n) iluminar la superación de un problema de la vida real. Es, sin duda, una visión positivista que asume que la realidad tiene una explicación y que los hechos suceden por una cadena de acontecimientos previos. Todo puede explicarse y la causalidad, en determinados momentos, puede auxiliar tal explicación de la realidad. En resumen, descubrir y verificar los supuestos, las cadenas causales, los compromisos que se esperan y, en general, toda la estructura de evolución de una teoría del cambio asimilada para atender una problemática, remite al equipo evaluador a la etapa de diseño del programa o política. Por lo general, siempre resultará conveniente suponer que el diseño es incorrecto y que las presunciones son incoherentes; ello, para verificar su solidez y coherencia. Para alcanzar el éxito, toda política o programa debe sustentarse en una teoría y debe solucionar problemas sociales y públicos, en lugar de supeditarse a caprichos del gobernante en turno, a compromisos electorales o, peor aún, a ocurrencias o corazonadas.

El cambio, el progreso y la mejora de las condiciones de vida de los ciudadanos, encarnan procesos exigentes, pues no son productos del azar ni de la buena suerte, mucho menos de la buena voluntad de los gobernantes. Exigen creatividad, rigor, soporte social e inteligencia, ingredientes no siempre presentes en el proceso de política pública; por tanto, la labor del evaluador siempre será fructífera y parte de la idea de que estos son escenarios poco probables. El cambio es un proceso intencional y solo se logrará en la medida en que su ruta esté bien cimentada.

REFERENCIAS

- Aguilar, L. F. (2000). *La implementación de las políticas*. Miguel Ángel Porrúa.
- Breuer, E.; Lee, L.; Silva, M. D. y Lund, C. (2016). Using theory of change to design and evaluate public health interventions: a systematic review. *Implementation Science*, 11: 63. <https://doi.org/10.1186/s13012-016-0422-6>
- Chevalier, J. M. y Buckles, D. J. (2009). SAS² *Guía para la investigación colaborativa y la movilización social*. <http://ecosad.org/laboratorio-virtual/phocadownloadpap/METODO-IMPLICATIVAS/sas2-chevallier-y-buckles-2009%201.pdf>
- Fisher, F. y Gottweis, H. (2012). *The Argumentative Turn Revisited. Public Policy as Communicative Practice*. Duke University Press.

- Gilbert, N. (ed.). (2001). *Researching Social Life*. Sage.
- Grant, L. V. (1973). Aspects of Theory Evaluation. *British Journal of Political Science*, 3(1), 99-111.
- Howlett, M. (2011). *Designing public policies: principles and instruments*. Routledge.
- James, C. (2011). *Theory of Change Review*. A report commissioned by Comic Relief. Comic Relief's Vision. https://cnxus.org/wp-content/uploads/2022/04/James_ToC.pdf
- Jordan, A. J. y Turnpenny, J. R. (eds.). (2015). *The Tools of Policy Formulation: Actors, Capacities, Venues and Effects*. New Horizons in Public Policy Series. Edward Elgar Publishing.
- Kellogg, W.K. Foundation (2004). *Using Logic Models to Bring Together Planning, Evaluation, and Action. Logic Model Development Guide*. Consultado el 12 de abril de 2019 de: www.bttop.org/sites/default/files/public/W.K.%20Kellogg%20LogicModel.pdf
- Leeuw, F. L. (2012). Theory-based evaluation and contribution analysis: Three problems and a few solutions. *Evaluation*, 18(3), 348-363, <https://doi.org/10.1177/1356389012452051>. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.841.1395&rep=rep1&ctype=pdf>
- Mayne, J. (2015). Useful Theory of Change Models. *Canadian Journal of Program Evaluation*, 30(2). <https://doi.org/10.3138/cjpe.230>
- Msila, V. y Setlhako, A. (2013). Evaluation of programs: Reading Carol Weiss, *Universal Journal of Educational Research* 1(4): 323-327. <https://doi.org/10.13189/ujer.2013.010408>.
- Neuman, W. L. (1997). *Social Research Methods. Qualitative and Quantitative Approaches*. Allyn and Bacon.
- Rossi, P. H.; Freeman, H. E. y Lipsey, M. W. (1998). *Evaluation. A Systematic Approach* (6a. ed.). Sage.
- Vesely, A. (2008). Problem Tree: A Problem Structuring Heuristic. *Central European Journal of Public Policy*, 2, 68-81.
- Vogel, I. (2012). Review of the use of Theory of Change. The UK Department for International Development (DFID). Consultado el 12 de abril de 2019 de: https://www.theoryofchange.org/pdf/DFID_ToC_Review_VogelV7.pdf
- Weiss, C. H. (1998). *Evaluation: methods for studying programs and policies*. Prentice Hall.
- Wholey, J. S.; Hatry, H. P. y Newcomer, K. E. (eds.). (2010). *Handbook of Practical Program Evaluation*.: Jossey-Bass A Wiley.

CAPÍTULO 6

La evaluación de la efectividad

INTRODUCCIÓN

El actuar de las instituciones y de sus agentes, así como el cumplimiento de sus tareas y responsabilidades es un foco adicional que, regularmente, se escruta, pues, al fin y al cabo, la superación de los problemas de vivir en sociedad en buena parte depende de ello. La evaluación del desempeño, bien de las personas, las organizaciones, las instituciones, de las políticas e incluso de las sociedades, es una dimensión en la que la evaluación suele concentrarse. Medir el desempeño es una preocupación permanente en las organizaciones debido a que están orientadas a la consecución de objetivos y metas y que, en cierto sentido, justifican su existencia. El desempeño está ligado, de manera directa, con el actuar de las agencias en la implementación de las políticas públicas y de las esferas gubernamentales, siendo un asunto sensible, toda vez que la prestación de los servicios públicos exige de buenas, si no las mejores, dosis de efectividad, lo cual es apenas obvio.

Por lo anterior, las personas en su diario vivir se ven afectadas por la efectividad de la prestación de dichos servicios: el abasto de agua, por ejemplo, su oportunidad y calidad determinan los niveles de salud. El servicio de transporte, la seguridad, etc., son asuntos que ponen a prueba el desempeño de las autoridades y las agencias gubernamentales. Indistintamente de nuestra condición, todos quisiéramos niveles notables de desempeño en el sector público, so pena de la insatisfacción y la protesta social. Entonces, las estructuras gubernamentales están bajo permanente observación y esta condición hace que la opinión pública sea dinámica, intensa y logre niveles de alarma y preocupación. En la era de la información y las telecomunicaciones es un asunto crucial mostrar niveles aceptables de efectividad.

Este capítulo está dedicado a la evaluación de la efectividad como un factor fundamental para la acción colectiva en el logro de sus metas o propósitos, abordando su concepto, determinantes, así como la gerencia como un proceso continuo y holístico. El foco central es el desempeño institucional y las técnicas administrativas desarrolladas para mejorar el actuar organizacional. Es de resaltar que la efectividad es uno de los principales atributos para el ejercicio del poder, pues los electores no viven de promesas y sus expectativas pueden llevar

al traste las carreras de los políticos de oficio. Sin resultados satisfactorios será difícil permanecer en el poder, a menos que sea por la fuerza. Sobra advertir que, en el quehacer público, el imperativo legal aconseja el cabal cumplimiento de las responsabilidades o, en caso contrario, asumir las sanciones establecidas.

6.1 SIGNIFICADO DE LA EFECTIVIDAD

De manera simple, la efectividad está vinculada con los logros alcanzados por una persona, organización, institución o sociedad con relación a sus metas o en el cumplimiento de las responsabilidades predeterminadas. El *Diccionario de la lengua española* (RAE, 2019) equipara la efectividad con aquella “capacidad de lograr el efecto que se desea o se espera”. Esto incluye, por supuesto, aquellos resultados alcanzados o cumplidos, los cuales son una contribución de los individuos, los grupos de trabajo, sus planes y políticas, sus condiciones contextuales y logísticas, etc. Asimismo, estos resultados comprometen muchas dimensiones de la actividad emprendida y van más allá de los logros económicos, mismos que son importantes, pero que las superan e, incluso, las opacan. La efectividad, el logro o el alcance de los resultados deben ser apreciados en una perspectiva más amplia, considerando otras cuestiones que comprometen lo social, lo ambiental, lo político, lo axiológico, etc., y que hacen que el éxito personal, organizacional o societal sea considerado holísticamente, es decir, de una manera integral (figura 6.1).



FIGURA 6.1. LA EVALUACIÓN DE LA EFECTIVIDAD

Fuente: elaboración propia.

Como fue mencionado en el capítulo inicial, la evaluación de la efectividad es un proceso que vincula la fase de planeación, aquel momento en que se determinan los objetivos, metas, estrategias y tácticas con los resultados (*outputs* y *outcomes*). Por lo regular, la evaluación se emprende al terminar la etapa de implementación, una función o los esfuerzos organizacionales, mediando distintos recursos y capacidades. En ese sentido, es por excelencia una evaluación *ex post*, lo que significa que los individuos deben realizar algunas actividades para cumplir con las metas y objetivos que deben desempeñarse.

El desempeño es el ejercicio de las obligaciones inherentes a una profesión, cargo u oficio. El *Oxford English Dictionary* (2020) define el *performance* como “cumplimiento, ejecución o realización de algo ordenado o emprendido”. El cumplimiento establece un vínculo con lo obtenido al final de un proceso y establece una correspondencia entre lo deseado y lo obtenido e indica también si la actividad fue hecha correctamente en la medida en que el propósito se alcanzó. En consecuencia, el acto de hacer algo bien o de manera exitosa, según el lenguaje cotidiano, es una clara referencia a lo que, por lo regular, se denomina efectividad. Desempeñarse de manera adecuada, dando cumplimiento a ciertas condiciones y superando aquellas dificultades emergentes, conduce al éxito de los emprendimientos. Las relaciones, aunque estrechas o imperceptibles, ahí están: el desempeño adecuado lleva a la efectividad.

En el sector público, el desempeño está estrechamente vinculado con el cumplimiento de la ley, pues no existe agencia u organización que no derive su actividad de un acto legal; en los actuales momentos, la legalidad es el origen de todas las cosas. Las distintas organizaciones del sector público son establecidas con responsabilidades, funciones y competencias precisas que deben cumplirse en apego a las disposiciones. Su sobrevivencia, legitimidad y reconocimiento pasan por la efectividad de su actuar. Aquellas agencias que no logren dichos estándares estarán condenadas a la desaparición, en especial cuando distintas voces demandan la mejora de la función pública.

6.2 ANTECEDENTES DE LA EFECTIVIDAD

El estudio de la efectividad no es reciente, pues los individuos, los gobernantes y los grupos siempre han perseguido objetivos. Los estudiosos de la administración, desde Henry Fayol, Frederick W. Taylor y sus sucesores, han estado preocupados por la efectividad. Hace poco, la gerencia del desempeño, función dedicada a estudiar la efectividad, aparece para superar las ausencias y debilidades de modelos y enfoques administrativos derivados de la administración por objetivos. Bajo esta perspectiva, el desempeño es dinámico, continuo y comprometido con la evaluación (Latham *et al.*, 2007, citado en Armstrong, 2009). Es una función permanente de los directivos y no un asunto exclusivo de las oficinas de recursos humanos.

La preocupación por la efectividad también se justifica por las nuevas condiciones de la economía del mercado que determina e impone los nuevos niveles de productividad y competitividad centrados en el cumplimiento sistemático de las metas organizacionales. Las organizaciones tienen el imperativo de hacer todo para conseguirlas; en caso contrario, morirán o desaparecerán. En momentos de recesión, escasez y alta competencia, el tema se vuelve relevante y determinante para la supervivencia. Desde una perspectiva evolutiva, solo sobreviven los que alcancen sus propósitos, los fuertes y dominantes. La preocupación por la efectividad compromete la capacidad de las organizaciones para cumplir sus responsabilidades y

atender los intereses de sus *stakeholders*, que esperan y demandan el éxito en desempeño organizacional.

El éxito es algo consustancial a la naturaleza humana. En ocasiones, se piensa que se vive para lograr el éxito, aunque este sea difícil de conceptualizar y acotar. En efecto, las aspiraciones personales, los sueños, las pasiones, las ambiciones y los anhelos, así como las responsabilidades de la vida, ejercen una presión constante para luchar, de manera constante, por el logro de propósitos. Esto aplica a gobiernos, organizaciones y sociedades en general.

En la política, la búsqueda del éxito es una obsesión y, sin temor a la equivocación, muy pocos buscan el fracaso. Aunque no se reconozca, Nicolás Maquiavelo (1469-1527), a través de su inmortal obra, *El príncipe*, fue de los primeros en estudiar la efectividad en la acción política. Allí, expone de manera sistemática los éxitos y fracasos de los más poderosos y propone algunas fórmulas para lograr, en el ejercicio del poder, el tan anhelado éxito. Desde su óptica, la victoria es lo más relevante y es la única medida del éxito, y alcanzar una meta es el único criterio para valorar las decisiones políticas; se es exitoso en la medida que se alcance una meta. La fórmula del éxito que propone está basada en dos condiciones determinantes: el azar y el obrar del gobernante. Mantener el Estado (*mantenere lo stato*) es el objetivo primordial de todo monarca y Maquiavelo recomienda utilizar aquellos métodos efectivos que se encuentren a su alcance. “El afán constante es anticiparse virtuosamente a la fortuna y derrotar a los adversarios” (Iturralde, 2015, p. 91). En ese sentido, el insigne florentino es un adelantado en el estudio de la efectividad y en la evaluación de las políticas emprendidas para regir los destinos de las ciudades-Estados. La verdad no difiere en mucho de los propósitos que agobian a los actuales gobernantes.

Por ello, en los tiempos actuales, la reforma del Estado y sus correspondientes olas o versiones son el motivador más poderoso de la evaluación en general y, de la efectividad, en lo particular. En efecto, el espíritu de las reformas administrativas en el sector público ha sido animado por ese afán de imprimir efectividad a la acción gubernamental y de mejorar el desempeño de las agencias. Las últimas décadas del siglo xx atestiguaron que muchos países optaron por varias acciones para elevar los niveles de calidad, efectividad, eficiencia y responsabilidad de los gobiernos, motivados por la pérdida de confianza y credibilidad ciudadanas. Fueron cambios deliberados a las estructuras, a las funciones y a los procesos de las organizaciones públicas para mejorar su funcionamiento y recuperar el beneplácito social. No se descartan, por supuesto, preocupaciones sobre la legitimidad de los gobiernos.

La ola reformista tuvo razones exógenas y endógenas. Desde el punto de vista académico se invoca a la Nueva Gestión Pública (*New Public Management*, NPM) como uno de los enfoques con intenciones reformistas de la gestión weberiana y basado en los resultados. La NPM describe el conjunto de técnicas del sector privado, aplicables al sector público para mejorar su nivel de desempeño. Implica también una transformación de las formas de gerenciar y gestionar los asuntos públicos, y de tramitar y lidiar con los mercados internos. Inicialmente en el Reino Unido, se trasladó después a Estados Unidos, Australia y Nueva Zelanda, pa-

ra luego impregnarse en la administración en los países escandinavos y la Europa continental. Considerado una revolución administrativa que impactó en distintos grados a todos los países (Lane, 2000). Por otra parte, Pollit y Bonckaert (2014) afirman que:

[...] en su primer nivel, la NPM puede considerarse como una teoría o doctrina general, que presupone que el desempeño del sector público puede mejorarse con la importación de conceptos, técnicas y valores de la administración de negocios. En su segundo nivel puede considerarse como un ramillete de conceptos y prácticas específicas, con marcado énfasis en el estudio del desempeño, la especialización de las organizaciones, la mediación de los contratos de desempeño, la inserción de los mecanismos de mercado; una tendencia a considerar a los usuarios como “clientes” y la aplicación de enfoques de calidad (2014, p. 10).

Lo cierto es que dicho enfoque puede considerarse un desafío a los modelos tradicionales o clásicos de la administración del sector público cuestionados y acusados de inoperantes frente a los desafíos actuales de las sociedades contemporáneas. Para otros constituye una nueva teoría para el sector público (Hood, 1991), un debate sobre los avances de la disciplina administrativa en los países avanzados en su afán por combinar la democracia con los esquemas del mercado para asignar recursos, (re)distribuir la renta y regular las actividades económicas (Lane, 2000). Asimismo, se trata de una modalidad gerencial para reconfigurar las funciones gubernamentales (interno), así como sus relaciones con el sector privado (externo). Es considerado también un nuevo paradigma o una revolución científica-gerencial para el gobierno del sector público, con miras a mejorar el desempeño de las agencias gubernamentales en ejercicio de sus funciones.

Otras perspectivas como la gestión del capital humano, la administración del recurso humano y otros enfoques, refuerzan la perspectiva del desempeño, toda vez que los logros organizacionales van más allá de los aportes individuales. Empezar a apreciar al funcionario como colaborador no como subordinado, cuyas contribuciones se deben gerenciar, potenciar, evaluar y retroalimentar. El desempeño individual contribuye, de manera sustancial, al desempeño organizacional.

6.3 DETERMINANTES DE LA EFECTIVIDAD

A menudo se reconoce que el desempeño es resultado de muchas condiciones o factores. Esto significa que es una variable dependiente que podría estar determinado por distintas cuestiones como aquellas inherentes al trabajador, o bien, a las relaciones que establezca o apoyo recibido y las condiciones institucionales (Harrison, 1997, en Armstrong, 2009). Las primeras tienen que ver con su entrenamiento, sus capacidades, motivación e incentivos que le estimulan; las segundas, con las formas como se relaciona el trabajador con su grupo, el tra-

bajo colaborativo que pueda generar y la influencia mutua para efectuar la tarea encomendada o con los referentes y apoyos derivados del jefe inmediato (estilo de dirección), y las últimas están relacionadas con la estructura, los valores o la cultura organizacionales. Como se aprecia, el desempeño depende de un cúmulo de condiciones a las que se encuentra sometido el responsable de las tareas (figura 6.2).

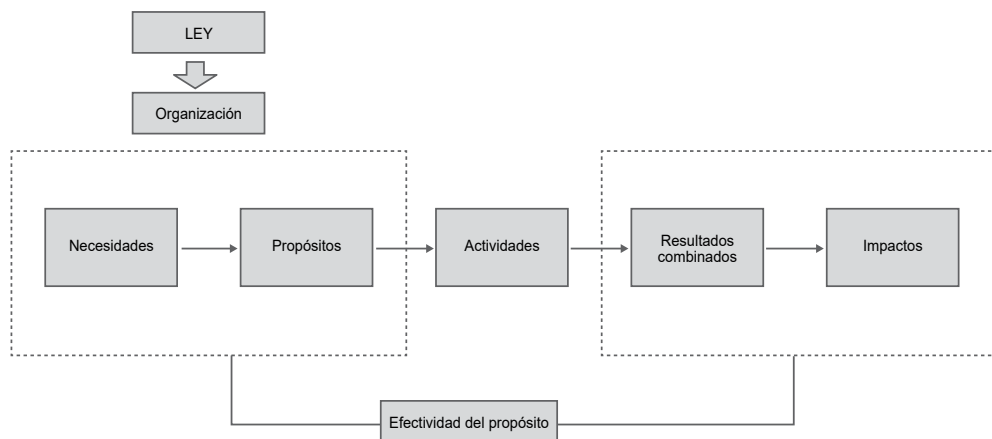


FIGURA 6.2. EL CIRCUITO DE LA EFECTIVIDAD

Fuente: elaboración propia con base en el gobierno australiano, Departamento de Finanzas, 2015.

Una manera sencilla para representar el desempeño en función de la interacción multiplicativa entre la habilidad y la motivación del empleado (Vroom, 1964, citado en Armstrong, 2009) es la siguiente:

$$\text{Desempeño} = f(\text{Habilidad} \times \text{Motivación})$$

Una propuesta más integral va más allá de lo individual, incorporando aportes del trabajo y de la organización, de modo que el desempeño es resultante de (Blumberg y Pringle, 1982, citado en Armstrong, 2009):

$$\text{Desempeño} = \text{Atributos individuales} \times \text{Esfuerzo de trabajo} \times \text{Apoyo organizacional}$$

Otros autores introducen la oportunidad del trabajador de participar en los procesos de trabajo (Boxall, Purcell y Wright, 2003, citado en Armstrong, 2009):

$$\text{Desempeño} = f(\text{Habilidad} + \text{Motivación} + \text{Oportunidades de participación})$$

De modo particular, la efectividad (figura 6.3) en el sector público está condicionada y asociada a las metas organizacionales y a su contexto (Walker, Boyne y Brewer, 2010). La ambigüedad es una de las particularidades del desempeño de actividades y asuntos de este sector. En los estudios de estos investigadores, la imprecisión domina las funciones, responsabilidades y competencias. Ello genera desconcierto, fomenta el choque institucional y promueve la confusión organizacional. En la formulación de las metas de cualquier organización pública, la vaguedad salta a la vista por cuenta de verbos imprecisos como propender por ...; fomentar el (la)...; propiciar... Expresiones estas muy comunes en planes de desarrollo o estratégicos, en leyes orgánicas y en manuales de procedimientos de agencias estatales. Al respecto, el enfoque estratégico enseña que son los verbos precisos de las formulaciones los que orientan la evaluación (incrementar, reducir, etcétera).

Por otra parte, los recursos limitados constituyen un generador importante de los bajos o precarios niveles de desempeño organizacional. Sin recursos suficientes, los servicios públicos y las funciones del sector público se ven diezmados en cobertura y calidad. Casi nunca se observa correspondencia entre los planes de operación y el respaldo financiero necesario, independientemente de que su cumplimiento sea prioritario para el gobierno en turno. Esto, a todas luces, obedece a un sobredimensionamiento de las promesas y a una limitada correspondencia presupuestal.

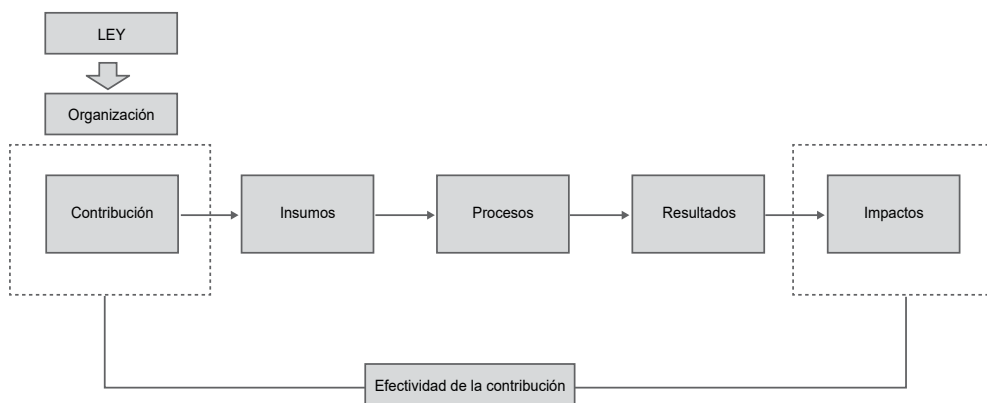


FIGURA 6.3. CONTRIBUCIÓN A LA EFECTIVIDAD

Fuente: elaboración propia con base en el gobierno australiano, Departamento de Finanzas, 2015.

Desde la teoría administrativa, las estructuras institucionales siempre se identifican como una condición para el cumplimiento de las funciones y competencias organizacionales; sin embargo, las investigaciones en el sector público destacan que no siempre se garantiza. Como es apenas lógico, este aspecto constituye una importante restricción para un adecuado desempeño y una cabal efectividad. Una de las cuestiones que, por lo normal, preocupa en el sector, gira en torno a quién diseña, con qué lógica, cuáles es su conocimiento y experien-

cia en organizaciones o agencias públicas. En este frente, pocas veces se garantiza que diseño, idoneidad y estabilidad institucional sean las características propicias para promover un desempeño idóneo. A esta disfuncionalidad se suma el hecho de que entidades incompetentes o disfuncionales permanezcan en el directorio, contribuyendo al desprestigio y desconfianza institucional.

A pesar de ser una condición necesaria, la coordinación institucional no se consigue siempre. Cuando la contribución a los propósitos institucionales es resultado de esfuerzos conjuntos o complementarios, dicha interdependencia limita la efectividad. La interdependencia, la complementariedad o la subsidiariedad organizacional constituyen obstáculos cuando la alta gerencia no logra el concurso de los esfuerzos o la concurrencia para la coproducción de los servicios públicos. En realidad, son las relaciones intergubernamentales las principales enemigas de la acción gubernamental y es que el postulado que los políticos ordenan y los burócratas obedecen, cada día pierde más vigencia.

Finalmente, la gerencia pública también atenta contra el buen desempeño y la efectividad. Lo estratégico se supedita a lo operativo y bajo esta jerarquía nada funciona, menos cuando la contratación de colaboradores no obedece a talentos, conocimientos, capacidades y habilidades, sino a la recomendación (o imposición) política. Con algunas excepciones, el acto de gobernar es un experimento, una cuestión de ensayo y error, y una aspiración pasajera a la que muchos apuestan.

6.4 LA GERENCIA DEL DESEMPEÑO

La literatura especializada en asuntos de evaluación invoca la evaluación del desempeño (*performance evaluation*), ello para referir la evaluación de la efectividad. De igual manera, desempeño hace referencia al proceso o a la manera de funcionar u operar exitosamente. Al interior de una organización, el desempeño se reconoce como complejo y como resultado de la combinación de muchas variables. Por ello, “el desempeño es un constructo multidimensional, una medida que varía de acuerdo con una serie de factores” (Bates y Holton, 1995, citado en Armstrong, 2009, p. 30). Latham, Sulsky y MacDonald (2007) enfatizaron la necesidad de una definición apropiada del desempeño como prerrequisito de la retroalimentación y del proceso de definición de metas. Sugieren entonces que un marco teórico para el desempeño debe considerar dimensiones, niveles y normas esperadas para distintos niveles, así como la identificación previa de sus restricciones, la manera como estas deben sopesarse y cómo debe fundamentarse, a fin de hacer comparaciones. Debe definirse por los resultados del trabajo, por dar cuenta del vínculo entre metas estratégicas organizacionales, por la satisfacción de públicos objetivos y contribuciones a la economía (Bernardin *et al.*, 1995, citado en Armstrong, 2009). Estas definiciones han permitido desarrollar perspectivas para estudiar y mejorar la efectividad organizacional.

La gerencia del desempeño es una función formalmente reconocida y con funciones concretas. Quienes intentan estudiar la dirección y evaluación del desempeño coinciden en definirlo como un proceso, una secuencia de actividades que van desde formulación de metas, establecimiento de normas, entrenamiento, retroalimentación, ajuste, compensación, etc., para obtener o mejorar los resultados de una persona, organización, institución o sociedad. A este tenor, para Armstrong (2009):

Es un proceso sistemático tendiente a la mejora del actuar organizacional, por medio del mejoramiento del desempeño individual y colectivo. Permite obtener resultados destacados de cara a una serie de metas, normas y requerimientos competitivos mediando un marco de entendimientos compartidos acerca de lo que debe ser alcanzado y sobre la forma en que deben gerenciarse las personas para incrementar la probabilidad de lograrlo a corto y a largo plazo (2009, p. 9).

Como proceso operativo, la gerencia del desempeño es sistemático al considerar diferentes dimensiones del actuar humano y grupal que impacta continuamente, la efectividad y la eficiencia, el compromiso y la actitud de los participantes en la consecución de las metas agregadas, superiores y colectivas. En Estados Unidos, la medición del desempeño obliga el monitoreo y el reporte del progreso para el cumplimiento de las metas preestablecidas de un programa, proyecto o política en una agencia o conjunto de agencias (*General Accountant Office*, GAO, 1998). Es un proceso que podría dar cuenta de los insumos, el proceso, los productos o los resultados de las actividades emprendidas. Se enfoca en determinar si se han alcanzado los objetivos expresados como una medida estándar del desempeño, manteniendo no solo el afán de tomar medidas administrativas para mejorar la prestación del servicio y la misión del programa, agencia o institución, sino también la de rendir cuentas.

El proceso considera distintos momentos (figura 6.4), partiendo del establecimiento de la misión institucional, justificando su existencia, determinando los valores que orientan el actuar y estableciendo con antelación los compromisos con sus públicos, para después establecer las metas para un periodo específico con sus respectivos indicadores de evaluación.

Desde la gerencia del sector privado también se considera el desempeño de la competencia (*benchmarking*), ya que permite a las organizaciones fijar sus propios derroteros en un ambiente competitivo. Los acuerdos o pactos de desempeño constituyen un momento especial para establecer los compromisos frente a las tareas estipuladas. Ello refleja una perfecta identificación de los involucrados frente a las acciones futuras, como también la necesidad y la importancia de los aportes individuales a los logros colectivos. Las directivas deben hacer un llamado al trabajo colaborativo, de otra manera, los planes y programas no se cumplirán. Prosigue la implementación o puesta en marcha de los planes y programas, el seguimiento continuo, la medición y la comparación de los resultados (con los previos o con los de la competencia). Este proceso de valoración se acompaña de recompensas y reconocimientos por el trabajo individual y colectivo.

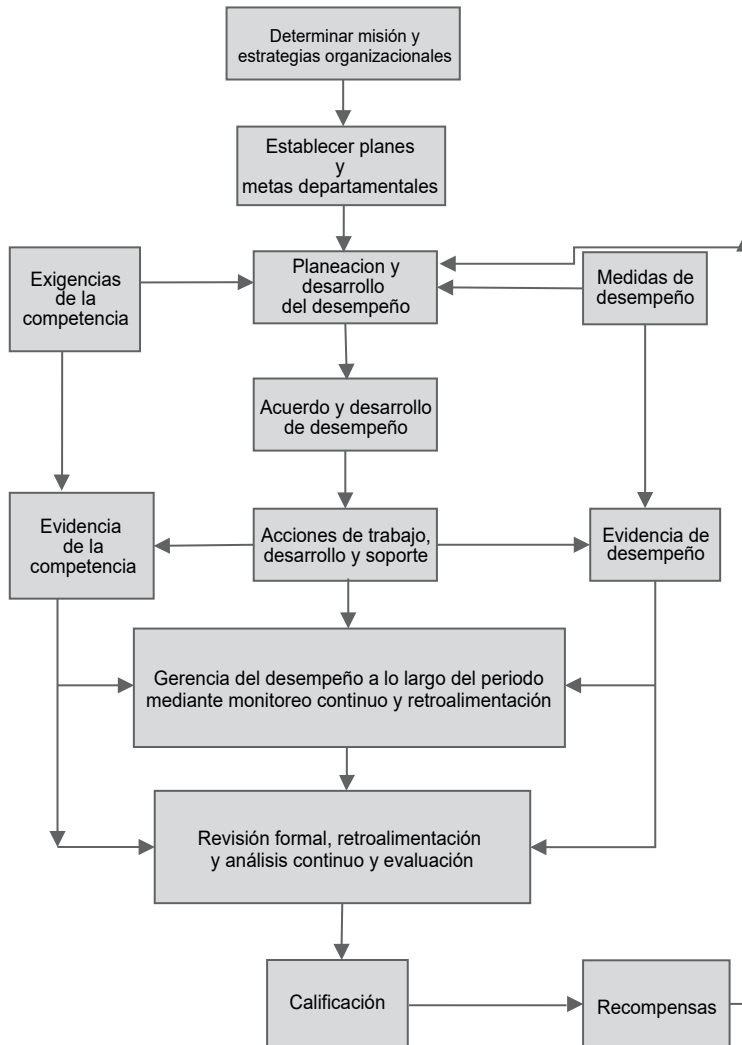


FIGURA 6.4. EL PROCESO DE LA GERENCIA DEL DESEMPEÑO

Fuente: Armstrong, 2009, p. 63.

6.4.1 Técnicas para investigar el desempeño

Dependiendo del tipo de organización, en la gerencia del desempeño se reconocen diferentes modalidades, focos y técnicas. De modo que puede recaer en el desempeño de los trabajadores, los insumos, los procesos, los resultados intermedios y/o los resultados finales. Asimismo, se dispone de distintos instrumentos y métodos, tales como gráficas, *rankings*, escalas, etc., que involucran distintos actores (internos y externos) o momentos.

6.4.2 *Desempeño basado en evidencias*

Los principios de la gerencia del desempeño en las prácticas organizacionales orientan el acopio de evidencias. Asimismo, significa la generación de información interna apoyada en la investigación y las ciencias sociales y su posterior análisis para inferir el desempeño real, identificando los factores que lo desvían y decidiendo rutas a seguir. Para ello es necesario establecer toda una estrategia de investigación que permita enfocarse en los asuntos críticos y recolectar la evidencia necesaria, reconociendo los problemas subyacentes a distintos niveles de desempeño observados.

Con anterioridad, es necesario definir algunos asuntos y establecer estrategias de medición, instrumentos de recolección y escalas de medición. Por tanto, también es necesario que todos en la organización (administrativos y empleados) acuerden asuntos básicos como: qué se considera un buen desempeño y cómo debe ser estudiado, medido y evaluado este. A la par, antes de adoptar este enfoque, es necesario establecer medidas referentes de la efectividad y entrenar los equipos en aspectos analíticos con técnicas apropiadas y robustas. Generar medidas del desempeño demanda inversiones en términos de esfuerzos organizacionales y de tiempo. Asimismo, requiere de equipos altamente calificados que asuman las actividades de manera rigurosa, sostenida y justa.

De igual manera, el enfoque basado en evidencias plantea una discusión acerca de los propósitos analíticos o evaluativos del mismo. Algunos prefieren los primeros, argumentando que en la gerencia del desempeño lo que realmente importa es el futuro, no el pasado y que, en consecuencia, se expresa mayor preocupación en mejorar los niveles alcanzados, que en juzgarlos (evaluativo). Con esta apreciación no concuerdan aquellos que basan el pago de acuerdo con lo ejecutado. En lo que la mayoría de los expertos sí coincide es que el análisis y la evaluación del desempeño basado en evidencias, es difícil y desafiante. Lograr una aceptación generalizada, generar la información, establecer estándares en torno a ellas, analizar y juzgar, encarna desafíos importantes que se deben asumir con la mayor seriedad, rigurosidad y constancia, so pena de generar frustración y malestar organizacional.

6.4.3 *Retroalimentación 360°*

Es una técnica que incorpora visiones sistemáticas de los *stakeholders* sobre el actuar de individuos y organizaciones. De esta manera, la evaluación se convierte en un asunto multidimensional e integral basado en dos supuestos: 1) cualquier discrepancia entre cómo se ve un empleado y cómo lo ven los demás, mejora la autopercepción; y 2) mejorar la autopercepción es fuente de mejoramiento del propio desempeño (Turnow, 1993, citado en Armstrong 2009). A través de la retroalimentación se conocen las debilidades¹ y fortalezas y pueden in-

¹ Conviene mencionar que ninguna escuela del pensamiento estratégico, ni los seguidores del estratega Igor Ansoff, han cambiado la palabra "debilidad" o "amenaza", por el concepto "área de oportunidad"; en la actualidad, la expresión se usa

corporarse visiones e intereses externos, así como fomentar la participación y el compromiso de todos los involucrados.

La metodología de la evaluación del desempeño 360° exige la preparación de instrumentos para el acopio de datos que miden la conducta de los individuos desde distintas perspectivas y que, después, se relacionará con las competencias. Se consideran, por tanto, asuntos de liderazgo, trabajo en equipo, comunicación, habilidades organizativas, resolución de conflictos, influencia y adaptabilidad. Las competencias y conductas se acompañan de escalas que denotan importancia, por ejemplo, las no importantes para el cargo y las esenciales. A la par, el desempeño de las primeras se asocia con lo débil y el de las últimas con lo destacado o sobresaliente de la conducta. Una vez acopiada esta, debe procesarse y retroalimentarse para después iniciar las correspondientes acciones.

Con relación a lo conveniente de la estrategia de 360° se identifican fortalezas, así como desventajas o debilidades. Dentro de las primeras destaca la generación de conciencia en los individuos acerca de cómo son percibidos por los demás, indistintamente de su nivel. Como ventaja se señala la visión global sobre el desempeño y la conciencia sobre la importancia de las competencias. En otras dimensiones, se resalta el descubrimiento de necesidades de desarrollo, las cuales deben ser apoyadas por los superiores. El proceso y la retroalimentación son percibidos como válidos y objetivos, y su aceptación es más probable. Por otra parte, como desventajas figuran los procesos burocráticos generados, los costos asociados en términos de tiempo y recursos y, en ocasiones, la falta de objetividad en las apreciaciones sobre la conducta de los individuos.

6.4.4 *Coaching*

Una vez conocidas las evaluaciones de las diferentes actividades, es necesario usar estos insumos para mejorar las competencias, habilidades, el conocimiento y, en consecuencia, el desempeño de los individuos. El *coaching* es un enfoque personal (usualmente uno-a-uno) que podría ayudar, ya que no se trata solo de una novedad, pues que cada vez que un gerente delega tareas, está generando una oportunidad de *coaching* en la medida que permite al colaborador aprender algo nuevo, así como considerar nuevas estrategias de trabajo, potenciar habilidades y competencias, y cumplir con la tarea encargada. “El *coaching* redefine las relaciones entre el supervisor y el subordinado. El trabajo conjunto ayuda al trabajador a tener un mejor desempeño” (Lee, 2005, citado en Armstrong, 2009, p. 167). Si bien, considera intervenciones cortas y momentáneas, privilegia el desempeño futuro autónomo y la mejora continua. El *coaching* debe considerarse un proceso formal bien concebido (planeado), pues no se trata de una supervisión ocasional o espontánea que aconseja, alienta o evalúa.

para “maquillar” los asuntos internos y externos que demandan actuación. Como siempre, los estrategias invitan a llamar las cosas por su nombre, ya que es la única manera de convocar esfuerzos en torno a planes de mejora.

Coachear consiste en establecer una relación directa, recolectar y analizar información, diagnosticar, retroalimentar las acciones planeadas, implementar las actividades previstas y volver a verificar el desempeño y los progresos (Thompson, Purdy y Summers, 2008, citado en Armstrong, 2009). Las etapas que considera el *coaching* son:

1. Descubrimiento por parte del individuo de su propia visión, de manera concomitante con la de otros, identificando fortalezas, debilidades, etcétera.
2. Plan de acción que incluye las metas y objetivos para desafiarse y desarrollarse en el contexto organizacional.
3. Revisión de logros a la luz de lo planeado (Woodruffe, 2008, citado en Armstrong, 2009). Obviamente, los planes de acción pueden ser reformulados y ajustados a la luz de lo acontecido. Quien emprende el *coaching* debe tener habilidades para escuchar, poseer cualidades analíticas y de indagación y desarrollar técnicas de retroalimentación que le permitan generar ambientes de mejoramiento permanente.

6.5 MEDIDAS DEL DESEMPEÑO

Las medidas del desempeño se han vuelto populares en el sector público. Smith (1990) advierte que la ubicuidad de los indicadores en el sector público es una condición y que hay una explosión en distintos niveles de la actividad gubernamental. En efecto, los sistemas de información en sus agencias son cada vez más comunes y los indicadores son omnipresentes. Al parecer, la gerencia del desempeño está inevitablemente sometida a la famosa sentencia, atribuida, en algunas ocasiones a Sir William Thomson, Lord Kelvin (siglo XIX) y, en otras, a Peter Drucker (1954, p. 114): “Si no se mide lo que se hace, no se puede controlar y si no se puede controlar, no se puede dirigir y, si no se puede dirigir, no se puede mejorar”. En efecto, el establecimiento de metas obliga a contar con un referente mensurable, a fin de verificar su cumplimiento a lo largo del tiempo.

En la evaluación de la efectividad se enfatiza la medición de resultados e impactos; sin embargo, a menudo se requiere la recolección de información de fuentes complementarias. Incluye, por tanto, datos directos de algunas actividades empleando entrevistas o haciendo el seguimiento de contactos o clientes en distintos momentos del programa. Las entrevistas con empleados complementarían la información requerida (Wholey, Hatry y Kathryn, 2010). Como cualquier proceso investigativo, el equipo evaluador debe procurar las fuentes de información y, sobre todo, su sistematización. Los sistemas de medidas del desempeño pueden ser pensados como herramientas de evaluación y/o de gerencia, diseñadas para proveer información valiosa, retroalimentar, fortalecer el proceso de toma de decisiones y mejorar el actuar de las organizaciones (Poister, 2010).

Los sistemas de información para estudiar el desempeño procuran información en varios sentidos. Inicialmente, tienen que ver con los resultados (*outputs*) que están representados por productos o servicios inmediatos logrados por las organizaciones. Estos pueden tener distintas expresiones, como unidades producidas durante un periodo o el número de cursos ofrecidos sobre algún tema específico. Lo que sí puede asegurarse es que los resultados inmediatos pueden lograrse de distintas formas y el evaluador tiene que desarrollar habilidad para generar información precisa y adecuada. En segundo lugar, deben dar cuenta de los resultados de más largo plazo o impactos (*outcomes*) que pudieran obtenerse con una intervención. Por ejemplo, contar con un sistema de monitoreo que establezca las medidas de aquellos impactos generados por un programa para reducir la incidencia de una enfermedad o de un programa de seguridad en las carreteras para disminuir la accidentalidad y las muertes. Qué decir de la calidad del servicio, de la satisfacción del ciudadano o del beneficiario. En muchas ocasiones, por razones de ignorancia o de apego ideológico, dichas dimensiones no se consideran como medidas del desempeño organizacional. No sobra advertir lo desafiante y costoso que pudieran ser las acciones de monitoreo y medición de resultados e impactos, pues al exigir un seguimiento sistemático una vez, el programa o intervención ha culminado. Adicional a que pudiera implicar largos periodos, ya que los impactos no pueden apreciarse de manera inmediata.

6.5.1 Las metas organizacionales

Todas las organizaciones tienen una razón de ser, uno o varios propósitos y, por supuesto, tienen metas que cumplir. En las organizaciones, los procesos de conformación, seguimiento, evaluación y reconfiguración de propósitos y metas son continuos. Al establecer metas, estas determinan las actividades o las acciones para alcanzarlas, pero cuáles son las razones para establecer metas. En primera instancia, incrementa la motivación, los individuos ajustan su comportamiento y sus actitudes a fin de alcanzarlas; encontrando así mecanismos de refuerzo y estrategias de motivación y estímulo del desempeño. En segundo lugar, las metas mejoran el desempeño y los individuos pueden enfocar sus habilidades y esfuerzos en la consecución de lo específico, de lo concreto. En tercer lugar, las metas elevan la eficacia de las personas al aumentar la persistencia al canalizar esfuerzos en su consecución (Turkay, 2014).

Más allá de los beneficios de su definición, las metas pueden ser procedimentales, al identificar estrategias para su logro, o bien, metas fijadas en función de resultados concretos. En la consecución de las metas se encuentran algunos factores determinantes del éxito como el entrenamiento, la motivación y otras condiciones contextuales que favorecen su consecución, como el apoyo técnico, financiero, etc., pasando por la autoeficiencia y el compromiso de las personas u organizaciones. Cuando se habla de metas (y de objetivos), por lo regular se invoca el acrónimo SMART para especificar sus atributos.

SMART es el acrónimo en inglés para condensar cinco cualidades: específicas (*specific*), medibles (*measurable*), alcanzables (*attainable*), relevantes (*relevant*) y oportunas (*timely*). Es-

pecíficas, con ello se enfatiza la necesidad de que las declaraciones y los propósitos sean claros, sin dar lugar a la ambigüedad o a la malinterpretación. Medibles, a fin de tener posibilidad de comparar a lo largo del tiempo o entre diferentes agentes. Alcanzables o que se puedan lograr, las inalcanzables garantizan el fracaso. Relevantes, en la medida que tengan importancia para los individuos, las organizaciones, las instituciones o las sociedades; la trivialidad no tiene cabida. Oportunas, en el momento y en la dirección indicada.

Las metas determinan la efectividad organizacional, esta mide qué tan apropiadas son aquellas y qué tan bien la organización está lográndolas. La efectividad es un atributo esbozado y perseguido de manera permanente, al punto que hoy constituye una de las características más anheladas. En la preocupación por la efectividad subyacen otras funciones de la administración, relacionadas con la dirección y el control. Las autoridades tienen la responsabilidad de establecer las metas, pero también de emprender acciones para su logro y control.

El desempeño de los individuos es mejor cuando tienen metas específicas y alcanzables (Latham y Locke, 1990, citado en Boxall, *et al.*, 2007). Una condición adicional al establecimiento de metas es el involucramiento de quienes las formulan, pues no deben ser impuestas, ya que la participación en el proceso incrementa su aceptación y con ello el desempeño individual. La participación en la planeación incrementa el compromiso, adicional a que su aceptación se refuerza cuando el individuo las percibe como justas y razonables, cuando cree en la administración y en el apoyo que recibirá, y cuando tienen certeza de que existirán las condiciones para lograrlas.

6.5.2 Tipos de metas

La primera referencia hacia las metas está vinculada con lo que se espera del trabajo, con la función primaria y con la naturaleza del cargo. Así, por ejemplo: “remitir todos los documentos recibidos a las distintas áreas de gestión” o “procesar la información semanalmente” establecen algunas metas que son deseables para un cargo particular. Cuando la meta se cuantifica de manera precisa (una cifra o dato y una fecha), algo siempre deseable, establecen los resultados medibles en algunas variables de la organización. Así, por ejemplo, reducir 50% de las órdenes reprocesadas durante el mes, representa una meta clara y específica. Por otra parte, podría hacerse referencia a metas a cumplir durante un tiempo y representar una tarea acumulativa para el final del periodo. Siguiendo con el ejemplo previo, sería diseñar una nueva línea de producción al término del año.

6.6 SÍNTESIS CAPITULAR

Medir el desempeño se ha convertido en una práctica aceptada, se ha sofisticado e institucionalizado tanto en el sector público como en las organizaciones no gubernamentales. Los sis-

temas de monitoreo del desempeño se diseñan para dar seguimiento, de manera regular, a ciertas medidas de los programas y reportarlas a las autoridades y mejorar así sus procesos de decisión y fomentar el *accountability* con los distintos *stakeholders*.

La evaluación del desempeño y el logro de la efectividad en lo personal y en lo organizacional no son asuntos menores ni propósitos banales, por el contrario, constituyen la parte más difícil de la misión institucional. Ello constituye una de las cuestiones que preocupa tanto a directores como a colaboradores, especialmente en épocas de creciente competitividad y demanda por el uso adecuado de los recursos.

Tales iniciativas deben ser sistemáticas y sostenidas con fuerte apoyo de la alta gerencia, a fin de generar una cultura de la mejora continua. Requiere planeación, análisis y retroalimentación. Siendo esta última, clave en la gerencia del desempeño y su evaluación, un asunto primordial para asegurar la efectividad del proceso. Los involucrados deben proveerse su propia retroalimentación y recibirla de otros, sus pares o sus supervisores. Es definitiva para ajustar el desempeño, visualizar el progreso en sus habilidades y tareas, y mejorar su actuación. Complementariamente, los reconocimientos por logros sobresalientes estimulan el desempeño.

Cuando los evaluadores se preguntan por qué algunas organizaciones e instituciones tienen éxito mientras que otras fallan, las posibles respuestas pueden encontrar razones en varias hipótesis endógenas sobre el desempeño y sobre la capacidad de las organizaciones. En tiempos recientes, otras variables exógenas han complementado las respuestas, al punto que los determinantes del desempeño se asocian al desarrollo socioeconómico, donde la cultura cívica es más activa y donde la estabilidad social es evidente (Putnam, 1983). Para el institucionalismo y desde la perspectiva del *path dependency*, la historia cuenta para el éxito. Como se aprecia, las cuestiones internas son importantes, pero también los asuntos contextuales.

REFERENCIAS

- Armstrong, M. (2009). *Armstrong's Handbook of Performance Management. An evidence-based guide to delivering high performance* (4a. edición). Kogan Page.
- Blumberg, M. y Pringle, C.D. (1982). The Missing Opportunity in Organizational Research. Some Implications for a Theory of Work Performance. *Academy of Management Review*, 7(4), 560-569.
- Boxall, P.; Purcell, J. y Wright, P. (2007). *The Oxford Handbook of Human Resources*. Oxford University Press.
- Drucker, P.F. (1954). *The Practice of Management*. Butterworth-Heinemann.
- General Accountant Office (1998). *Performance Measurement and Evaluation. Definitions and Relations*.
- Hood, C. (1991). A Public Management for all Seasons? *Public Administration*, 69, 3-19.
- Iturralde, I. (2015). *De príncipes, caciques y otros animales políticos*. Boanalletra Alcompas.

- Lane, J.E. (2000). *New Public Management*. Routledge.
- Latham, G.; Sulsky, L.M. y Macdonald, H. (2007). Performance management, 364-381, en: Peter Boxall, John Purcell y Patrick Wright (eds.), *Oxford Handbook of Human Resource Management*. Oxford University Press.
- Poister, T. H. (2010). Performance Measurement Monitoring Program Outcomes, 100-124, en: Joseph S. Wholey, Harry P. Hatry y Kathryn E. Newcomer (eds.), *Handbook of Practical Program Evaluation*, (3a. edición). Jossey-Bass books.
- Real Academia Española. *Diccionario de la lengua española* (23ª edición). Consultado el 12 de octubre de 2019. <https://dle.rae.es>
- Smith, P. (1990). The Use of Performance Indicators in the Public Sector. *Journal of the Royal Statistical Society*, 153(1), 53-72.
- Turkay, S. (2014). Setting Goals: Who, Why, How? Manuscript. <https://vpal.harvard.edu/publications/setting-goals-who-why-how>
- United States, General Accounting Office (2005). *Performance measurement and evaluation. Definitions and Relationships*. (Supersedes GAO-05-739SP). <https://www.gao.gov/assets/a202676.html>
- Walker, R. M.; Boyne, G. A. y Brewer, G. A. (eds.) (2010). *Public Management and Performance. Research Directions*. Cambridge University Press.
- Wholey, J. S.; Hatry, H. P. y Newcomer, K. E. (2010). *Handbook of Practical Program Evaluation* (3a. edición). Jossey-Bass.

CAPÍTULO 7

Evaluación de la eficiencia económica

INTRODUCCIÓN

Para los fines del estudio de la eficiencia económica y bajo el supuesto de este libro, una política pública puede asimilarse a un flujo de dinero, o lo que es lo mismo, como entradas y salidas de recursos intentando alcanzar resultados a corto, mediano y largo plazos. En esta dirección, la evaluación de la eficiencia toma sentido e intenta determinar qué tan bien se han utilizado dichos dineros. La eficiencia es una cuestión que preocupa no solo al asignador del gasto, sino también a sus fuentes y financiadores. La ingeniería económica permite estudiar, de manera cuidadosa, los movimientos del dinero a la luz de ciertas variables financieras y la garantía de la eficiencia. Este capítulo presenta algunos temas relacionados con la eficiencia económica y su evaluación, utilizando indicadores como valor presente neto, relación beneficio/costo y el indicador de costo-eficiencia.

Es imposible negar que la evaluación económica en el sector público y, en particular, de las políticas públicas, genera controversia en la medida que algunos argumentan que, en dicho sector, la eficiencia no debería ser motivo de preocupación. A pesar de esto, la evaluación de la eficiencia ya es de interés creciente, al punto que es condición inamovible para muchos financiadores como donantes. Por ello, es necesario que los conceptos y los principales indicadores de eficiencia sean de pleno dominio entre los involucrados, a fin de comprender, aplicar e interpretar su significado. En consecuencia, este capítulo en sus inicios presenta las reflexiones acerca de la eficiencia para, posteriormente, desarrollar todos los fundamentos de la ingeniería económica y sentar las bases para la construcción de indicadores, a fin de concepcuar sobre la eficiencia y procurar elementos o criterios para la toma de decisiones.

7.1 LA EFICIENCIA ECONÓMICA

La consideración de eficiencia económica para las actividades del sector público genera muchas resistencias, pero lo cierto es que, nadie argumenta o aclara la improcedencia de su apli-

cación. La eficiencia es un atributo, parámetro o criterio estudiado por la economía y su referencia es la eficiencia propuesta a finales del siglo XIX, por Vilfredo Pareto. La primera idea sobre la eficiencia en el uso de los recursos públicos es aquella que se considera con o sin proyecto. La eficiencia en la asignación de recursos para una política sería juzgada en términos de si ella fuera relativamente superior *-pareto óptima-*, con relación a la situación en ausencia de dicha política. Otra idea respecto al uso eficiente de los recursos postula que, si como mínimo una persona mejora su bienestar y ninguna lo desmejora, se alcanza la eficiencia. O bien, si quienes se benefician de una política logran compensar a aquellos que resultan perjudicados (ninguno empeora), la política pública promoverá la optimalidad paretiana (Fuguitt y Wilcox, 1999). Desde otra perspectiva más simple, la eficiencia se suscita cuando los beneficios generados por la política, programa o proyecto son mayores que los costos ocasionados (la sociedad como un todo). Ello, sin considerar quién recibe los beneficios, quién asume los costos o si es o no es compensado por las posibles pérdidas. En este tipo de consideraciones en torno a la eficiencia, lo más importante es la mejora del bienestar de algunos de los involucrados, sin desmedro del bienestar de otros. En ello existe consenso de que el gasto público tiene este tipo de potencialidades y responsabilidades y posibilita también una de las funciones más importantes para el sector público en procura del bienestar y el progreso de las personas.

La figura 7.1 ilustra la situación de la eficiencia en sentido de Pareto. Se trata de observar los intercambios entre dos productos dada una frontera de posibilidades, fijada por la cantidad de recursos disponibles como factores de producción.

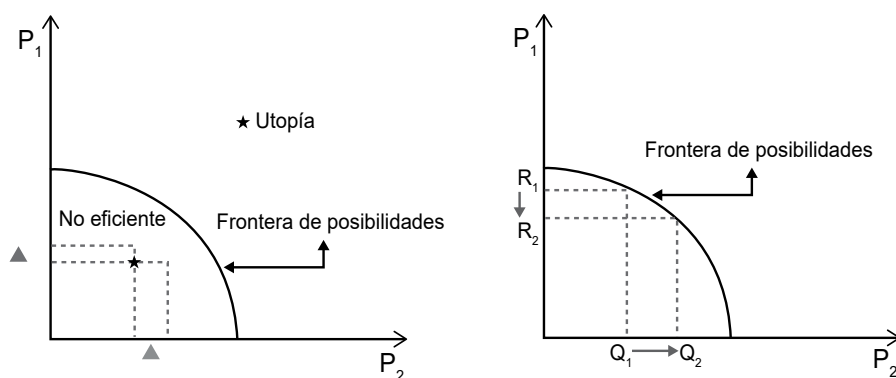


FIGURA 7.1 EFICIENCIA PARETIANA

Fuente: elaboración propia con base en Stiglitz, 2000.

Existen tres áreas de análisis o tres posibles escenarios, según la utilización de los recursos disponibles. Al interior de la frontera se observa una situación donde no se alcanza dicha eficiencia, puesto que al incrementar las cantidades de un producto se observa que la cantidad

del otro producto no se afecta; incluso, se podría aumentar la cantidad de los dos simultáneamente. Desde la perspectiva de los fallos del mercado se dice que:

Las asignaciones de recursos que tienen la propiedad de que no es posible mejorar el bienestar de ninguna persona sin empeorar el de alguna otra, se dice que son eficientes en el sentido de Pareto u óptimas en el sentido de Pareto (Stiglitz, 2000, p. 69).

Pensar más allá de la frontera de posibilidades resulta utópico, dado que los recursos son finitos y están acotados por la frontera de posibilidades. La eficiencia solo puede ser alcanzada en la frontera de las posibilidades, pues si se quiere aumentar la cantidad del P_2 , pasando de la cantidad Q_1 a Q_2 , de inmediato la cantidad del producto P_1 disminuiría reduciendo su cantidad de R_1 a R_2 . La eficiencia paretiana indica que para alcanzar una mejora del bienestar se debe garantizar la mejora de alguien, sin desmedro de los niveles de bienestar de los otros.

7.2 CONCEPTOS BÁSICOS DE INGENIERÍA ECONÓMICA

La ingeniería económica estudia el comportamiento del dinero a lo largo del tiempo, integrando conocimientos financieros y microeconómicos. Su pretensión es apoyar la decisión y elección de inversión entre distintas alternativas disponibles, una vez identificados todos sus requerimientos y estimado los costos e ingresos durante el ciclo de vida. Desde una perspectiva comparada, esto es, entre alternativas el analista puede identificar la opción más favorable o determinar cuál de ellas es la más óptima desde el punto de vista económico.

El análisis económico de una política o de un proyecto, por lo regular, suele acompañarse de una representación gráfica que refleja, en una línea del tiempo, el ciclo de vida desde su concepción y expansión hasta su extinción. Es reconocido como el diagrama de flujo que muestra los flujos de dinero (entradas y salidas) ocasionados en cada periodo de la vida de un proyecto, programa o política. El momento cero de esta línea de tiempo es el presente y, cada proporción, la unidad de tiempo (año, mes, trimestre, etc.), indica los momentos posteriores para su ejecución y operación. En el diagrama, la dirección de las flechas hacia arriba indica la entrada de dinero o flujo positivo y una flecha hacia abajo, la salida de dinero o flujo negativo.

La figura 7.2 ilustra un diagrama de flujo de una política cuya duración o ciclo de vida es de cinco periodos, representados desde el momento cero al momento cinco. En el momento cero, el presente aparece una salida (desembolso) de \$500, representado por una flecha hacia abajo. Posteriormente, se presentan otras salidas (costos) periódicas de \$50, a partir del primer año y hasta el quinto. Al final, en el periodo 55 se presenta una flecha hacia arriba, lo que indica un ingreso (entrada) de \$800. P indica presente y F futuro (valores que ocurren una vez en un periodo único); A representa una serie uniforme de cinco salidas ($n =$

5) de dinero de \$50. Para que A se considere una serie, el monto del flujo debe ser uniforme y extenderse a lo largo de periodos consecutivos (\$50 en puntos 1, 2, 3, 4 y 5).

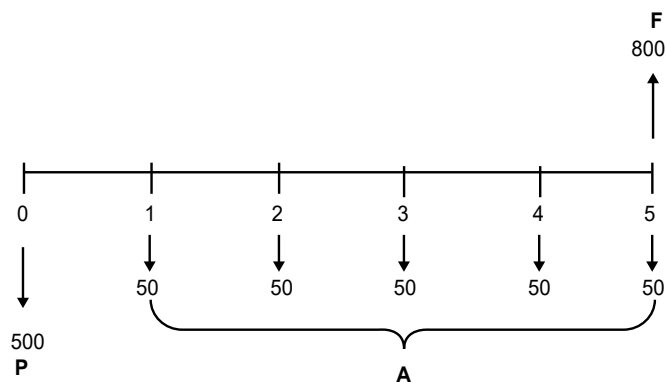


FIGURA 7.2. DIAGRAMA DE FLUJO DE UNA POLÍTICA PÚBLICA

Fuente: elaboración propia.

7.2.1 Principios del dinero

Como ya se advirtió, el objeto de la ingeniería económica es el dinero, por lo que es fundamental invocar los dos principios que, no solo regulan el comportamiento de sus distintos flujos, sino también se relacionan con las tasas de interés (conocimiento financiero) y el diseño de indicadores económicos.

Valor en el tiempo

En los estudios económicos suele afirmarse que “el dinero crea dinero”; este precepto resulta cierto, ya que si una persona decide invertir hoy (banco, bolsa de valores, cooperativa de ahorro, etc.), a futuro tendrá más dinero acumulado que el depositado inicialmente. A esta acumulación “se le denomina valor del dinero en el tiempo y constituye el concepto más importante de la ingeniería económica” (Blank y Tarquin, 2006, p. 9). Esta manifestación del valor del dinero en el tiempo se reconoce como interés, que es una medida del aumento entre la cantidad de dinero original y la final acumulada. La siguiente ecuación representa este crecimiento.

$$\text{Monto del interés} = \text{Cantidad total acumulada} - \text{Cantidad inicial}$$

Cuando el interés se expresa como un porcentaje de la cantidad original por unidad de tiempo, se obtiene una tasa de interés:

$$\text{Tasa de interés} = \frac{\text{Interés acumulado por unidad de tiempo}}{\text{Cantidad inicial}} \times 100$$

Aplicación práctica

Una compañía depositó en el momento uno, \$100.000 y, un año después, retiró \$106.000. ¿Cuál fue el monto del interés y cuál su tasa de interés?

$$\text{Monto del interés} = 106,000 - 100,000 = 6,000$$

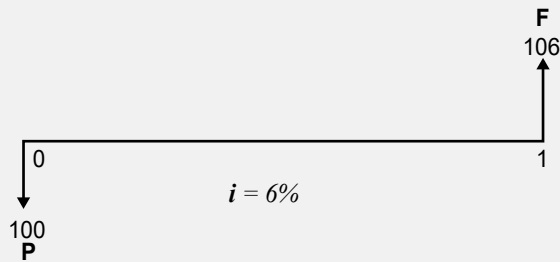
$$\text{Tasa de interés} = \frac{6,000}{100,000} \times 100 = 6\%$$

Equivalencia

La aplicación simultánea del valor del dinero en el tiempo y de la tasa de interés genera la equivalencia. Esto significa que sumas diferentes de dinero en distintos momentos de tiempo pueden ser iguales en valor económico.

Aplicación práctica

Si la tasa de interés durante un periodo es de 6%, \$100 de hoy equivaldrán a \$106 dentro de un periodo, es decir, si alguien le ofrece hoy \$100 o \$106 dentro de un año, no habrá diferencia. Por tanto, las dos sumas de dinero son equivalentes en valor económico, a pesar de que ocurren en distintos momentos (mayores detalles en Blank y Tarquin, 2006, p. 17).



7.2.2 Valor presente y valor futuro

De los dos principios previos puede derivarse un valor futuro a partir de un valor presente, conociendo la tasa de interés compuesto (i) prevaleciente y el número de periodos. Al suponer que existe un incremento en la cantidad de dinero que fue originalmente invertida (principal), expresada por el interés, la manera de calcular la cantidad futura después de transcurrir un periodo sería:

$$\text{Cantidad futura} = \text{Principal} + \text{Interés}$$

$$F_1 = \text{Principal} + \text{Principal} (\text{tasa de interés})$$

$$F_1 = P + P(i)$$

$$F_1 = P(1 + i)$$

Si esta cantidad futura calculada se somete de nueva cuenta a otro periodo, que se afectaría por la tasa de interés, se habla entonces del interés compuesto, al considerar la cantidad total de interés acumulado en periodos anteriores. Para el segundo periodo se tiene:

$$F_2 = F_1 + F_1 (i)$$

$$F_2 = P(1 + i) + P(1 + i)i$$

$$F_2 = P(1 + i) + P(1 + i)^2$$

$$F_2 = P(1 + i + i + i^2)$$

$$F_2 = P(1 + 2i + i^2)$$

$$F_2 = P(1 + i)^2$$

Asimismo, la cantidad de dinero acumulado al final del periodo tres sería:

$$F_3 = F_2 + F_2 (i)$$

$$F_3 = P(1 + i)^2 + P(1 + i)^2 (i)$$

$$F_3 = [P(1 + i) + P(1 + i)i] + [P(1 + i) + P(1 + i)i]$$

$$F_3 = [P(1 + i) + P(1 + i)i] + [P(1 + i)i + P(1 + i)i^2]$$

$$F_3 = [P(1 + i) + 2P(1 + i)i + P(1 + i)i^2 + i^2]$$

Factorizando $P(1 + i)$:

$$F_3 = P(1 + i)(1 + 2i + i^2)$$

$$F_3 = P(1 + i)(1 + i^2)$$

$$F_3 = P(1 + i)i^3$$

De los valores anteriores puede generalizarse la fórmula para n años como:

$$F_n = P(1 + i)^n$$

Que representa el pago futuro de una inversión inicial P , después de n años a una tasa de interés i .

Para el cálculo de un valor presente, dado un valor futuro se parte de la misma ecuación y para ello se despeja P en términos de F y se obtiene:

$$P = F \left[\frac{1}{(1 + i)^n} \right]$$

La expresión anterior permite determinar el valor presente P , dada una cantidad futura F cuando han transcurrido n periodos, a una tasa de interés i . Es importante anotar que las dos fórmulas son de pago único, es decir, se utilizan para encontrar cantidades presentes o futuras cuando se tiene un solo pago o entrada.

7.2.3 Serie uniforme

Como se había anticipado, una serie uniforme (A) es un flujo semejante en el monto, durante periodos continuos, siendo posible relacionarla y encontrarle su equivalente con un valor presente (P). Se trata de encontrar una equivalencia de un valor presente, a partir de una serie uniforme y comenzando al final del periodo uno y extendiéndose durante n años, a una determinada tasa de interés i . La derivación de la fórmula conduce a la siguiente ecuación:

$$P = A \left[\frac{(1 + i)^n - 1}{i (1 + i)^n} \right]$$

La expresión entre corchetes es conocida con la etiqueta "Factor Valor Presente Serie Uniforme (FVPSU)". Este factor dará el valor presente P de una serie uniforme equivalente A " (Blank y Tarquin, 2006, p. 60), que comienza al final del año uno y se extiende durante n

periodos a una tasa de interés i . Al reordenar la ecuación anterior, A puede expresarse en términos de P :

$$A = P \left[\frac{i (1 + i)^n}{(1 + i)^n - 1} \right]$$

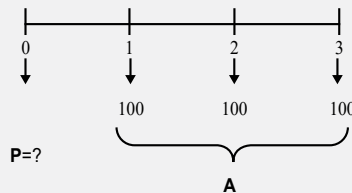
El término entre corchetes se denomina Factor de Recuperación de Capital (FRC), o factor A/P . Con él se calcula el valor anual uniforme equivalente A durante n años de una P dada en el año 0, cuando la tasa de interés es i (Blank y Tarquin, 2006).

Es importante recordar que las fórmulas se derivaron del valor presente P y con el primer valor del costo uniforme A , con un año de diferencia, “es decir, el valor presente P siempre debe localizarse un periodo antes de la primera A ” (Blank y Tarquin, 2006, p. 60).

Aplicación práctica

Suponga que una política pública tiene unos costos anuales constantes de \$100 durante sus primeros tres años, a una tasa de interés de 12%. Calcular el valor actualizado de la serie.

Primero se traza el diagrama de flujo con la información proporcionada y que exige determinar el valor presente de una serie uniforme; por tanto, en la ecuación P será la incógnita.



$$P = A \left[\frac{(1 + i)^n - 1}{i (1 + i)^n} \right]$$

$$P = 100 \left[\frac{(1 + 0.12)^3 - 1}{0.12 (1 + 0.12)^3} \right]$$

$$P = 100 \left[\frac{(1.12)^3 - 1}{0.12 (1.12)^3} \right]$$

$$P = 100 \left[\frac{1.4049 - 1}{0.12 (1.4049)} \right]$$

$$P = A \left[\frac{0.4049}{0.68588} \right] = 2.4018$$

$$P = 100 \times 2.4018 = 240.18$$

Lo anterior significa que una serie de \$100 durante tres periodos consecutivos a partir del año uno, en valor presente, equivale a \$240.18, cuando la tasa de interés es de 12%.

Alternativamente, el cálculo de esta equivalencia podría efectuarse considerando cada salida de dinero, no como una serie sino como un F (con $n = 1, 2$ y 3) y cada uno actualizado a una tasa de interés de 12%. Para tal efecto, debe utilizarse una ecuación que involucre a P y F , donde P es la incógnita. Veamos su desarrollo:

Se parte de la ecuación:

$$F_n = P(1 + i)^n$$

Se despeja P :

$$P = \frac{F}{(1 + i)^n}$$

$$P = (P/F, n = 1, i = 12\%) + (P/F, n = 2, i = 12\%) + (P/F, n = 3, i = 12\%)$$

$$P = \frac{100}{(1 + 0.12)^1} + \frac{100}{(1 + 0.12)^2} + \frac{100}{(1 + 0.12)^3}$$

$$P = \frac{100}{(1.12)^1} + \frac{100}{(1.12)^2} + \frac{100}{(1.12)^3}$$

$$P = \frac{100}{1.12} + \frac{100}{1.2524} + \frac{100}{1.4049}$$

$$P = 89.2857 + 79.7193 + 71.1794$$

$$P = 240.18$$

Con estos fundamentos de ingeniería económica, el equipo evaluador puede emprender el estudio de la eficiencia de una política, a partir de la construcción de indicadores que apoyan el análisis costo-beneficio.

7.3 ANÁLISIS COSTO-BENEFICIO

Los antecedentes de la evaluación identifican al ingeniero y matemático francés Jules Dupuit (1844) como el pionero del análisis costo-beneficio (AC/B), cuando estudió la utilidad de la inversión en obras públicas de su tiempo. Con frecuencia, el AC/B se somete al escrutinio, a tal grado que muchos se preguntan por el motivo de las controversias. Frank y otros (2000) lo descalifican como método para soportar decisiones sabias e inteligentes en el ámbito gubernamental. Otros, por su parte, como Richardson (2000) ponderan su utilidad buscando justificaciones y aplicaciones en distintos frentes de la vida social (Sen, 2000; Kornhauser, 2000; Becker, 2000; Posner, 2000; Broome, 2000). Sin embargo, lo cierto es que el enfoque no ha sido sustituido y permanece como una opción que, aunque imperfecta, proporciona algunos criterios para el análisis y la reflexión sobre la eficiencia económica. El AC/B soporta la toma de decisiones en el ámbito económico.

El AC/B implica la recolección y organización de información relevante para decisiones de líderes gubernamentales que intervienen, cuando los mercados fallan, mediante proyectos, programas o regímenes regulatorios públicos (Krutilla, 1961; Musgrave, 1969, en Miller y Robbins, 2007). Asimismo, constituye una investigación evaluativa centrada en decidir si se continúa o no con un programa, estrategia o política, o en asignar recursos entre programas en competencia (GAO, 1991; Poister, 1978; Weiss, 1972, en Miller y Robbins, 2007). Es un enfoque que evalúa si las decisiones o las elecciones que afectan el uso de recursos escasos promueven la eficiencia. Considerando una política específica y las alternativas relevantes, este análisis identifica sistemáticamente las consecuencias y evalúa los rendimientos y erogaciones sociales, aplicando criterios de decisión. Para ello, sitúa los costos y beneficios en el momento actual, en valor presente y los descuenta a una tasa de interés determinada. Por lo regular, la proporción de descuento es una tasa de interés tomada de los mercados financieros (Frank, 2000). Es un proceso de agregación o ponderación, tomando rendimientos y erogaciones para evaluar lo adecuado o eficiente de un proyecto (Broome, 2000).

Como cualquier otro ejercicio, la evaluación de la eficiencia exige el diseño de indicadores y el establecimiento de criterios de decisión, para que el decisor garantice cierta coherencia en la toma de decisiones. El valor presente neto, la tasa interna de retorno, la relación beneficio-costos y el costo-eficiencia se presentan como indicadores que sirven como base para el estudio de la eficiencia en iniciativas económicas, mismos que pueden llevarse al plano de la inversión gubernamental.

7.3.1 Valor Presente Neto (VPN)

El primer indicador económico para considerar en la evaluación económica es el Valor Presente Neto (VPN). Como su nombre lo indica, corresponde a la diferencia entre el valor actualizado de los beneficios y el valor actualizado de los costos descontados a una determinada tasa de interés i :

$$VPN = VP_b - VP_c$$

El VPN proporciona unos criterios de decisión que facilitan dicho proceso al inversionista, con tres criterios de decisión:

Resultado		Decisión
Positivo	$VPN > 0$	Se acepta
Nulo	$VPN = 0$	Indiferente
Negativo	$VPN < 0$	Se rechaza

Si el valor del VPN es mayor que cero, significa que los beneficios son mayores que los costos y, en virtud de ello, la decisión sería positiva, es decir, se acepta la alternativa. Si el VPN es igual a cero, significa que los beneficios son iguales a los costos y, en consecuencia, sería indiferente aceptar o rechazar la alternativa de inversión o deberían considerarse otros criterios de análisis. Si, por el contrario, el VPN es negativo, significa que los costos son mayores que los beneficios y, por tanto, la alternativa no es favorable desde el punto de vista económico. Cuando existen dos alternativas con VPN positivo e igual duración, se decide por el mayor VPN. Si existen dos alternativas con VPN negativo y debe llevarse a cabo el proyecto, se optaría por aquella con menor valor.

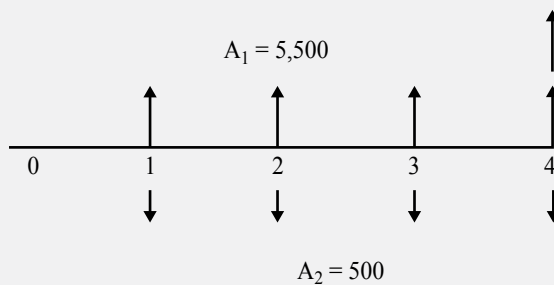
Aplicación práctica

“Una papelería está considerando la posibilidad de cambiar su fotocopiadora actual por una más moderna que tiene costos de operación menores y una mayor velocidad” (Vélez, 2011, p. 171). La vida útil de la máquina es de cuatro años, los costos y los ingresos de cada alternativa se presentan a continuación:

Ingresos y egresos	Fotocopiadora actual	Fotocopiadora nueva
Valor actual de venta	10,000	0
Valor actual de compra	0	16,000
Costo anual de operación	500	200
Ingresos anuales	5,500	8,500
Valor residual	7,000	9,000

La decisión de comprar o no una máquina nueva se tomará determinando el VPN calculado para cada una de las alternativas disponibles, no sin antes trazar el respectivo diagrama de flujo.

Actual

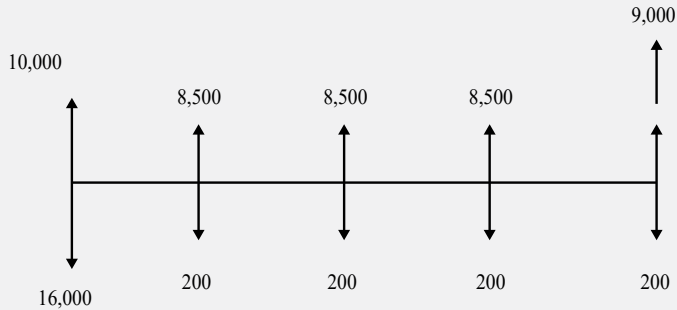


$$VPN_i = 5,500(3.0373) + 7,000(0.6355) - 500(3.0373)$$

$$VPN_i = 16,705.15 + 4,448.5 - 1,518.65$$

$$VPN_i = 21,153.65 - 1,518.65$$

$$VPN_i = 19,635$$

Nueva

$$VPN_2 = 10,000 + 8,500(3.0373) + 9,000(0.6355) - [16,000 + 200(0.0373)]$$

$$VPN_2 = 10,000 + 25,817.05 + 5,719.5 - 16,607.46$$

$$VPN_2 = 41,536.55 - 16,607.46$$

$$VPN_2 = 24,929.09$$

A la luz de los criterios de decisión se debe optar por la compra de la fotocopiadora nueva.

7.3.2 Relación Beneficio–Costo (B/C)

Es un indicador que permite establecer una relación entre los beneficios y costos en el momento cero, el presente. Es decir, encontrar la relación entre los beneficios y costos, descontados a una determinada tasa de interés. Una vez determinada dicha relación, se podrán aplicar los siguientes criterios de decisión:

Resultado		Decisión
Mayor que uno	$B/C > I$	Se acepta
Igual a uno	$B/C = I$	Indiferente
Menor que uno	$B/C < I$	Se rechaza

Si el cociente es superior a uno (I), significa que los beneficios son mayores que los costos y, en consecuencia, desde el punto de vista económico se aceptarían la alternativa, el proyecto

o la política. Si la relación es igual a uno (1), quiere decir que el nivel de costos es igual al nivel de beneficio y sería indiferente hacerlo o no. Finalmente, si la razón es inferior a uno (1), implica que los costos son mayores que los beneficios y debe, por tanto, rechazarse la alternativa desde lo económico.

Aplicación práctica

Al considerar, por ejemplo, los datos presentados en el ejercicio de las fotocopadoras, el evaluador podría realizar el cálculo de la relación B/C para cada situación: con proyecto y sin él.

Fotocopiadora actual

El valor actual de los beneficios es de \$21,153.65 y sus costos actualizados son de \$1,518.65, así que la relación B/C es la resultante de:

$$\frac{B_A}{C_A} = \frac{21,153.65}{1,518.65} = 13.92$$

Como es superior a uno, el criterio de decisión indica que es conveniente aceptar la actual fotocopiadora y por cada unidad de costo se generan \$13.92 de beneficio. Algo atractivo sin lugar a dudas.

Fotocopiadora nueva

$$\frac{B_N}{C_N} = \frac{41,536.55}{16,607.46} = 2.50$$

Para el caso en mención, la relación B/C de la fotocopiadora es igualmente positiva. Por cada unidad de costo se generan \$2.5 de beneficio, lo cual también es conveniente; sin embargo, dicha razón es muy superior en la fotocopiadora actual. En todo caso, el evaluador debería incorporar criterios adicionales para salir de esta situación tan particular. ¿Acaso es mejor una actualización tecnológica? ¿La velocidad de trabajo permite ampliar los actuales mercados?

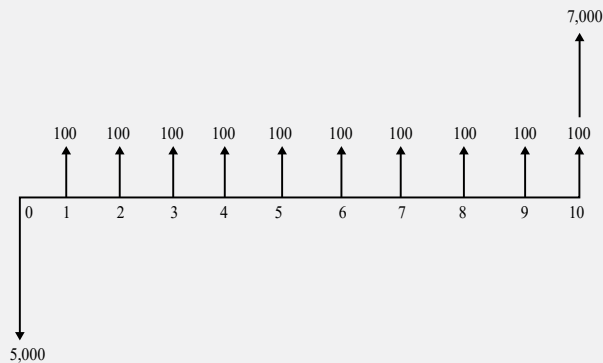
7.3.3 Tasa Interna de Retorno (TIR)

Un tercer indicador es la TIR, aquella tasa que iguala a cero el valor presente neto, es decir, que los beneficios actualizados son iguales a los costos actualizados. Mide la rentabilidad interna del proyecto y será pertinente compararla con tasas alternativas de inversión (tasas de oportunidad). Desde el punto de vista técnico, la TIR presenta dificultades para su cálculo

y, en ocasiones, genera tasas negativas, algo de naturaleza paradójica y, por ello, su uso no es tan popular.

Aplicación práctica

Si hoy se invierten \$5,000 en acciones y se espera que produzcan \$100 anuales durante 10 años y \$7,000 al final de 10 años ($n = 10$), se pide determinar la tasa interna de retorno. El diagrama de flujo se presenta a continuación.



Para calcular la TIR, se debe encontrar la tasa de interés que haga el Valor Presente Neto igual a cero. Es importante recordar que el Valor Presente Neto es aquel indicador que se obtiene restando los beneficios y costos actualizados y que tiene tres posibilidades: mayor que, igual a o menor que cero. La TIR es aquella tasa de interés que hace que el VPN sea cero, por lo que se tendría la siguiente relación:

$$0 = -5,000 + 100 \left(\frac{P}{A}, i, 10 \right) + 7,000 (P/F, i, 10)$$

$$P = \$5,000$$

$F = 10(100) + 7,000 = 8,000$ (Un cálculo sin tener en consideración la tasa de interés)

$$5,000 = 8,000(P/F, i, 10)$$

$$\frac{5,000}{8,000} = (P/F, i, 10)$$

$0.625 = (P/F, i, 10) \Rightarrow$ Dado un futuro encontrar un presente con i de incógnita.

Para calcular la TIR, el siguiente paso es remitirse a las tablas de interés compuesto y encontrar la tasa de interés que comprende el factor 0.625 cuando $n = 10$ y se conoce un valor futuro para determinar un presente (P/F). Efectivamente, se tiene que ir a la columna, dado un P/F con $n = 10$, buscando un límite superior y uno inferior que consideren el factor de la referencia 0.625. Para la tasa de 4%, el factor de valor futuro, dado un presente (P/F) con diez periodos, es 0.6756 y para la tasa de 5%, dado un P/F con $n = 10$, el factor es 0.6139. Quiere decir que la tasa que se está buscando es superior a 5% e inferior a 6%. Al aplicar dichas tasas se tiene:

$$0 = -5,000 + 100 \left(\frac{P}{A}, 5\%, 10 \right) + 7,000 (P/F, 5\%, 10)$$

$$0 = -5,000 + 100 \left[\frac{(1.05)^{10} - 1}{0.05(1.05)^{10}} \right] + \frac{7,000}{(1.05)^{10}}$$

$$0 = -5,000 + 100(7.7217) + 7,000(0.6139)$$

$$0 = -5,000 + 772.17 + 4,297.3$$

$$0 = -\$69.47$$

Ahora se aplicará la tasa de 6%, así:

$$0 = -5,000 + 100 \left(\frac{P}{A}, 6\%, 10 \right) + 7,000 (P/F, 6\%, 10)$$

$$0 = -5,000 + 100 \left[\frac{(1.06)^{10} - 1}{0.06(1.06)^{10}} \right] + \frac{7,000}{(1.06)^{10}}$$

$$0 = -5,000 + 100(7.3601) + 7,000(0.5584)$$

$$0 = -5,000 + 736.01 + 3,908.8$$

$$0 = \$355.19$$

Como se tiene un límite positivo y otro negativo, se procede a realizar la interpolación. Mediante este procedimiento se encuentra la tasa y el interés que hacen que el Valor Presente sea nulo o igual a cero. Esto es, que el valor presente de los ingresos sea igual al valor presente de los costos y, al restarlos, se iguala a cero. Para tal efecto se utilizará la siguiente convención:

$$-69.46 \quad 5\%$$

$$7 \quad i$$

$$355.19 \quad 6\%$$

$$c = a(d) = (69.46 - 0)(6 - 5)$$

$$b = 69.46 - (-355.19)$$

$$= 0.16$$

$$i = 5.00 + 0.16 = 5.16$$

Esta tasa debe generar un valor presente igual a cero. Veamos,

$$0 = -5,000 + 100 \left(\frac{p}{A}, 5.16\%, 10 \right) + 7,000 (P/F, 5.16\%, 10)$$

$$0 = -5,000 + 100 \left[\frac{(1.0516)^{10} - 1}{0.0516(1.0516)^{10}} \right] + \frac{7,000}{(1.0516)^{10}}$$

$$0 = -5,000 + 100(7.6623) + 7,000(0.6046)$$

$$0 = -5,000 + 766.23 + 4,232.42$$

En efecto, la tasa encontrada iguala el VNP a cero.

7.3.4 Indicador de Costo-Eficiencia (C/E)

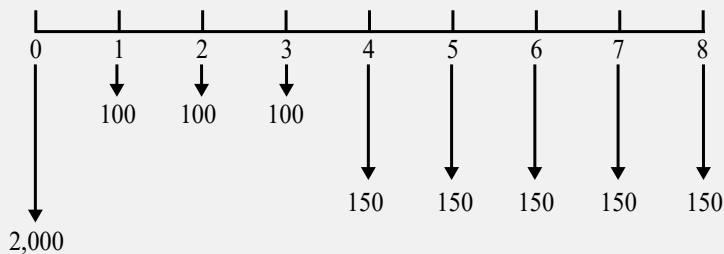
Cuando es posible cuantificar los beneficios de un proyecto o programa (números, cantidades, volúmenes, etc.) es posible calcular los indicadores de Costo/Eficiencia (C/E). Para esto, se utiliza en primera instancia el Costo Uniforme Equivalente (CUE), y después se vinculan al denominador, las unidades de beneficio o de producto que generará el proyecto (en promedio) durante su horizonte de operación. Este indicador es de gran importancia para conocer cuál es el valor de cada unidad de beneficio, cuestión que preocupa a los tomadores de decisiones o sobre el cual se fija un criterio de decisión.

Aplicación práctica

Supóngase un programa que tiene preestablecidos los siguientes costos, considerando una tasa de descuento de 12%:

- Inversión (momento cero) = \$2,000
- Costos anuales de operación del primer al tercer año = \$100
- Costos anuales de operación del cuarto al octavo año = \$150

El flujo de caja se representa de la siguiente manera:



Una forma usual de resolver esta situación es:

1. Traer todos los componentes a valor presente (al momento cero), descontados a la respectiva tasa de interés. Ya estando allí todos los valores pueden sumarse obteniendo un valor presente de los costos.
2. Hacer una distribución en ocho valores iguales (serie uniforme). Esto es, dado un valor presente encontrar un A , con $n = 8$. En la anotación de la ingeniería económica se tendría $(A/P, n = 8, i = 0.08)$. De esta manera se obtiene el CUE.

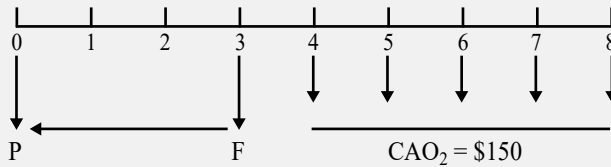
Regresando al ejercicio planteado, el primer paso requiere encontrar un VP de todos los flujos de dinero. El primer valor es la inversión que, al estar en el momento cero, constituye un valor presente y no requiere ninguna operación. De manera simple, \$2,000 de hoy equivalen a \$2,000. Complementariamente, los costos anuales de operación (CAO) se presentan por medio de dos series uniformes. La primera, denominada CAO_1 , son tres costos de operación cada uno de \$100 y se ocasionan del año uno al tres. La segunda serie, CAO_2 , tiene un valor de \$150 en cada año y va desde el periodo cuatro al ocho. Dichos valores deben actualizarse y, dadas unas series uniformes, determinarse los valores presentes. Según las convenciones utilizadas en ingeniería económica (P/A), con sus respectivos n a la tasa de interés de referencia. El valor presente del flujo de dinero es:

$$VP = \$2,000 + CAO_1 + CAO_2$$

$$VP = \$2,000 + (P/A, \$100, n = 3, 12\%) + (P/A, \$150, n = 5, 12\%)$$

Es importante advertir que la serie de costo anual de operación dos, CAO_2 , debe resolverse en dos pasos:

1. Convertir la serie en un valor presente que se ubica un año antes de la primera cuota de la serie, es decir, en el tercer año.
2. Trasladar ese valor (futuro) al año cero, con tres periodos.



$$VP = 150(\text{Factor de } A \text{ a } P, n = 5, i = 12\%) = 150(3.6048) = 540.72$$

$$P = 540.72 (\text{Factor de } F \text{ a } P, n = 3, i = 12\%) = 540.72(0.7118) = \$384.88$$

El valor presente total es resultante de los tres componentes:

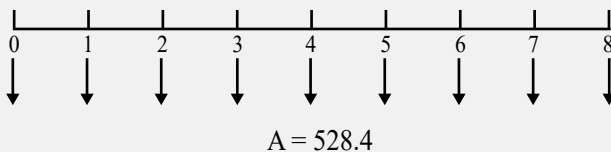
$$VP = \$2,000 + \$240.18 + \$384.88 = \$2,625$$

El segundo momento, una vez calculado el VP, consiste en encontrar el CUE, el cual solo demanda encontrar una serie uniforme, A , en este caso para ocho periodos.

$$CAE = VP(\text{Factor de } P \text{ a } A, n = 8; i = 12\%)$$

$$CAE = \$2,625(0.2013) = \$528.4$$

El perfil equivalente del proyecto sería:



Una vez calculado el CUE, puede incorporarse el beneficio obtenido durante cada periodo, a fin de calcular el indicador de costo-eficiencia. Supongamos que el programa cuyo perfil se describió tiene previsto una cobertura de 30 usuarios en cada periodo.

El indicador de costo-eficiencia medido por el CUE/Usuarios, resulta de la siguiente relación:

$$\frac{C}{E} = \frac{\$528.4}{30 \text{ usuarios}} = \$17.6/\text{usuario}$$

Este indicador representa el costo de cada unidad de beneficio y sirve para hacer comparaciones con iniciativas semejantes, ya sea al interior de una región, país e incluso a nivel internacional. Para las agencias de desarrollo, por ejemplo, interesa saber cuánto cuesta dar agua o luz a una vivienda. También es importante saber cuánto cuesta prevenir una enfermedad infantil con una vacunación o ilustrar a un estudiante. En general, este indicador es de especial importancia para quien debe tomar decisiones o para el ordenador del gasto, puesto que compromete la eficiencia en el uso de los recursos.

7.4 EVALUACIÓN ECONÓMICA EN EL SECTOR PÚBLICO

Por más que quisiera, el sector público no puede escapar al estudio de la eficiencia en el uso de los recursos. En sus dominios siempre rondará la preocupación de cómo utilizar, de mejor manera, los siempre escasos recursos. Los servicios públicos provistos siempre tienen una relación con estas preocupaciones, ya sea en la provisión de agua, en los servicios educativos o de la salud. Sin embargo, los gobiernos no son famosos por su eficiencia (Olson, 1973) y su desempeño siempre ha sido motivo de controversia. Aun así, sus representantes no son muy afectos a incorporar técnicas que les indiquen cómo han utilizado los recursos para cumplir con sus responsabilidades, ¿qué es lo que les disgusta?, ¿son sus técnicas o las conclusiones de su aplicación? Lo cierto es que, en algunos países, existe una renuencia marcada a la evaluación de la eficiencia, en especial con el análisis costo-eficiencia.

Las primeras experiencias del uso de tales principios datan de la década de los años treinta del siglo XX, cuando el gobierno federal de los Estados Unidos, mediante el Acta de Control de 1936, autorizaba solo aquellos proyectos de mejoramiento de la navegación y de los de control de inundaciones, cuyos beneficios excedían a sus costos. El Análisis Costo-beneficio se institucionalizó y rápidamente se volvió una técnica de análisis del sector hídrico en este país (Fuchs y Anderson, 1987). Desde entonces, dichas técnicas han colonizado otros sectores, como la defensa en tiempos de Robert McNamara como parte del famoso Sistema Presupuestal de Planeación y Programación (*Planning Programming Budgeting System*, PPBS), que después fue adoptado por el presidente Lyndon Johnson para la rama ejecutiva. A pesar del controvertido uso y de la resistencia de algunos sectores, el ACB se ha institucionalizado gracias a la aparición de nuevos enfoques de estudio de las actividades regulativas del gobierno y del interés de algunas ramas de la economía.

En América Latina existen algunas experiencias que indican que el uso de las técnicas del análisis costo-beneficio y sus derivados enfoques está siendo aceptados en algunas instancias gubernamentales. Las metodologías desarrolladas por el Banco de Proyectos de Inversión Nacional del Departamento Nacional de Planeación, en Colombia, representan un notable ejemplo y gozan de una trayectoria importante. En efecto, desde 1989 se dio comienzo a una iniciativa para fomentar una cultura de los proyectos en el sector público y para realizar evaluaciones *ex ante*. Ubicado al más alto nivel gubernamental, dicha modalidad ha venido fomentando las metodologías para identificar, formular y evaluar proyectos de inversión que aspiran a la financiación de los dineros públicos.

7.5 SÍNTESIS CAPITULAR

Con los principios de la ingeniería económica se facilita la construcción de indicadores para soportar la toma de decisiones. Al determinar los costos totales de cada alternativa en la fase de formulación, la información de carácter monetario se ordena y sistematiza para convertirla en insumo de la evaluación económica. La pretensión es obtener elementos de juicio para decidir en torno a la conveniencia de ejecutar un proyecto y sobre las formas eficientes de adelantarlos. Por tanto, son necesarias dos condiciones: ciertas alternativas y un objetivo.

Los promotores de proyectos y programas y los servidores públicos deben conocer, al detalle, las técnicas para tomar decisiones y determinar los procedimientos para la construcción de indicadores que ilustren el nivel de eficiencia obtenido. Si bien es cierto que los gobiernos no deben preocuparse en exceso por la eficiencia para la toma de decisiones, es necesario establecer instancias que analicen los costos antes y después de la implementación de sus políticas. El Banco de Proyectos de Inversión Nacional establecido en Colombia es un instrumento que alienta la cultura de los proyectos y fomenta el uso adecuado de los recursos públicos, asuntos que, sin duda, mejoran el desempeño del sector público.

REFERENCIAS

- Becker, G.S. (2000). A Comment on the Conference on Cost-Benefit Analysis. *The Journal of Legal Studies*, 29(S2), 1149-1152.
- Blank, L. y Tarquin, A. (2006). *Ingeniería económica*. Javier Brito (trad.) (6a. edición). McGraw-Hill Interamericana Editores.
- Broome, J. (2000). Cost-Benefit Analysis and Population. *The Journal of Legal Studies*, 29(S2), 953-970.
- Frank, R. H. (2000). Why is cost-benefit analysis so controversial? *Journal of Legal Studies*, 29(S2), pp. 913-930.

- Fuchs, E. P. y Anderson, J. E. (1987). The Institutionalization of Cost-Benefit Analysis. *Public Productivity Review*, 10(4), 25-33.
- Fugitt, D. y Wilcox, S. J. (1999). *Cost-Benefit Analysis for Public Sector Decision' Makers*. Quorum books.
- Kornhauser, L. A. (2000). On Justifying Cost-Benefit Analysis. *The Journal of Legal Studies*, 29(S2), 1037-1057.
- Krutilla, K. (2005). Using the Kaldor-Hicks Tableau Format for Cost-Benefit Analysis and Policy Evaluation. *Journal of Policy Analysis and Management*, 24(4), 864-875.
- Miller, G. J. y Robbins, D. (2007). Cost-Benefit Analysis, pp. 465-480, en: *Handbook of Public Policy Analysis. Theory, Politics, and Methods*. Frank Fischer, Gerald J. Miller y Mara S. Sidney (eds.). CRC Press.
- Olson, M. (1973). Evaluating Performance in the Public Sector, en: *The Measurement of Economic and Social Performance*, 355-409, <http://www.nber.org/chapters/c3618>
- Posner, R. A. (2000). Cost-Benefit Analysis: Definition, Justification, and Comment on Conference Papers. *The Journal of Legal Studies*, 29(S2), 1153-1177.
- Richardson, H. S. (2000). The Stupidity of the Cost-Benefit Standard. *The Journal of Legal Studies*, 29(S2), 971-1003.
- Sen, A. (2000). The discipline of cost-benefit analysis. *The Journal of Legal Studies*, 29(S2), 931-952.
- Stiglitz, J. E. (2000). *La economía del sector público* (3a. ed.). Antoni Bosch.
- Vélez, L. D. (2011). Aspectos económicos de los proyectos del sector público. Documento de trabajo. Disponible en: <https://webcache.googleusercontent.com/search?q=cache:i4LU4K26bcJ:https://repositorio.unal.edu.co/handle/unal/20057&cd=1&hl=es&ct=clnk&gl=mx>

CAPÍTULO 8

Evaluación económica del programa de la tortilla¹

INTRODUCCIÓN

Una vez sentadas las bases de la evaluación de la eficiencia, este capítulo presenta los resultados de la evaluación económica del Programa de Subsidio al Consumo de la Tortilla. Asimismo, constituye una experiencia real adelantada por el gobierno federal mexicano con una importante cobertura geográfica y social. Inicialmente, se presentan las reglas de operación del programa y, a continuación, se analizan los componentes del presupuesto destinado para atender los compromisos legales y, a partir de este, se construyen los indicadores necesarios. Para tal efecto, se emprende un análisis de costo-eficiencia en el cual se comparan los costos asociados a la inversión y operación del programa con los beneficios obtenidos durante 18 meses, a partir de enero de 2000.

Como podrá apreciarse, se trata de la evaluación económica de un programa social, soportada en indicadores de costo-eficiencia que apoyarán las acciones administrativas y los ajustes institucionales necesarios, en procura de un mejor desempeño de las agencias gubernamentales comprometidas. Se trata de una evaluación *ex post*, debido a que el programa llevaba operando varios años y lo que se pretendía era determinar el costo de ofrecer un kilo de tortilla subsidiado, con el propósito de que las autoridades competentes decidieran sobre el futuro del programa.

8.1 LINEAMIENTOS DEL PROGRAMA

En noviembre de 1990, la Compañía Nacional de Subsistencias Populares (CONASUPO) emprendió el Programa de Subsidio al Consumo de la Tortilla, bajo la modalidad de subsidio selectivo y dirigido. Por las dinámicas propias del sector público, el programa fue entregado primero a través del Fideicomiso para la Liquidación del Subsidio a la Tortilla (Fidelist) y, luego, a partir del año 2000, a Liconsas, S.A. de C.V. El programa buscaba apoyar a las fa-

¹ Este capítulo está basado en la evaluación del Programa de la Tortilla realizada para el Congreso de la Federación en el año 2001.

milias en pobreza extrema de las colonias marginadas de localidades urbanas de 15 mil habitantes y más, proporcionándoles un kilogramo diario de tortilla enriquecida, adicionada y fortificada. Para su entrega, se apoyaba en los proveedores de la industria de la masa y la tortilla del país. Para simplificar y reducir costos operativos (Acuerdo 2002 sobre las reglas de operación de la Secretaría de Desarrollo Social) se tuvieron que sistematizar sus procesos.

El programa adoptó ciertos criterios como objetividad, equidad, transparencia, selectividad, temporalidad y enfoque de género para la entrega del kilogramo de tortilla subsidiado. Para la elección de los beneficiarios se tuvieron en cuenta, entre otros asuntos: la situación demográfica, las características de la vivienda, las condiciones sanitarias, el nivel de ingreso, la posesión de enseres, particularidades socioeconómicas, nivel educativo y acceso a la seguridad social. Esta información fue originada en el Instituto Nacional de Estadística, Geografía e Informática (INEGI) y en el Consejo Nacional de Población (CONAPO). Desde luego, se identificaron los industriales de la masa y la tortilla (operadores del programa), considerados como el vehículo para entregar el subsidio a los beneficiarios. El programa estableció criterios de factibilidad y elegibilidad tanto para beneficiarios como para industriales, a fin de garantizar una adecuada focalización y un efectivo abasto del subsidio.

Un lector de banda magnética de la tarjeta del beneficiario permitía acceder al subsidio y, a la vez, enviaba información a Liconsa para el respectivo seguimiento y control diario. Para efectuar la evaluación pretendida es necesario analizar los asuntos del presupuesto público, la población a atender, así como los costos de operación y administración, para después construir los indicadores para conocer los niveles de eficiencia en la entrega de un kilogramo de tortilla en las 852 localidades del país. La población de industriales de la masa y la tortilla, operadores del programa o terceros participantes, se concentró, de manera significativa, en el Estado de México con 8.47%, seguido de Veracruz con 6.93% y Jalisco con 6.38% (mapa 8.1).



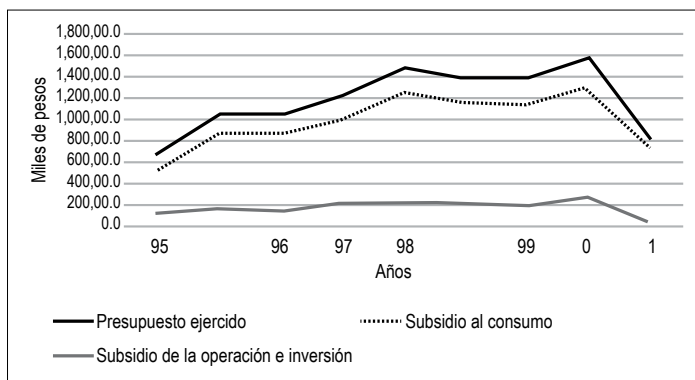
MAPA 8.1. INDUSTRIALES DEL PROGRAMA DE APOYO AL CONSUMO DE LA TORTILLA (NIVEL ESTATAL)

Fuente: elaboración propia con datos del Fidelist-Liconsa.

8.2 PRESUPUESTO DEL PROGRAMA

Por tratarse de un programa del ejecutivo federal, enmarcado en los programas para superar la pobreza, como lo estipula el Plan Nacional de Desarrollo 1995-2000, los recursos asignados eran de naturaleza pública y su ejercicio y control debían cumplir con lo estipulado en el Decreto de Presupuesto de Egresos de la Federación. Un análisis de dichos recursos, desde 1995, permite observar que su asignación se incrementó año tras año, en pesos corrientes, tanto para el subsidio pretendido como para el soporte de los gastos de operación e inversión. Esta situación permaneció hasta el año 2000, a partir de dicho año se redujeron significativamente los recursos (gráfica 8.1). Es de advertir que, en sus orígenes, los recursos del programa provenían del gobierno federal, pero en 2000 se suman los recursos propios del Fidelist, producto de sus propias actividades que vinieron a apoyar el programa. Aunque en proporción menor estos recursos deben señalarse como un recurso importante para apoyar los respectivos subsidios.

Por lo regular, la dinámica del presupuesto público revela dos momentos bien diferenciados: programación presupuestal y apropiación efectiva. Pero de igual modo, pueden darse diferencias durante una vigencia fiscal entre lo apropiado y lo ejecutado en realidad. Es posible afirmar que el programa no sufrió recortes importantes que pudieran menoscabar su desempeño durante el periodo de estudio. En efecto, considerando el presupuesto a precios constantes de 1995, se verifica un crecimiento sostenido de los recursos del programa hasta 1998; sin embargo, en 1999 presenta un descenso cercano a 16%, notándose de nueva cuenta un incremento posterior (gráfica 8.1).



GRÁFICA 8.1. PRESUPUESTO DEL PROGRAMA DE SUBSIDIO AL CONSUMO DE LA TORTILLA 1995-2000 (MILES DE PESOS)

Fuente: elaboración propia con datos del Fidelist-Liconsu.

Para efectos del análisis es importante aclarar que los dineros apropiados tenían dos destinos particulares: el pago a los industriales de la masa y la tortilla por el beneficio entregado (un

kilo de tortilla diario) a las familias en pobreza extrema, y los gastos de operación e inversión de la institución encargada de la administración del programa, tanto a nivel central como en las respectivas gerencias estatales. Esta estructura presupuestal indica que la mayor proporción de los recursos asignados eran entregados al subsidio al consumo (entre 80% y 88%), mientras que el restante correspondía al subsidio para la operación e inversión, cuyas proporciones se mantienen entre 1995 y 2001 (tabla 8.1).

Una posibilidad inicial para el análisis del programa se encuentra en el diseño de un indicador del nivel de ejecución de los recursos disponibles al relacionar lo asignado y lo ejecutado.

$$I_e = \frac{\text{Presupuesto ejecutado}}{\text{Presupuesto asignado}}$$

Donde:

I_e = Indicador de ejecución

$P.E.$ = Presupuesto ejecutado

$P.A.$ = Presupuesto asignado

TABLA 8.1. PRESUPUESTO GENERAL DEL PROGRAMA DE LA TORTILLA 1995-2001
(MILES DE PESOS)

Año	P. programado	P. asignado	P. ejecutado	%
1995		769,198.0	670,056.5	87.1
1996		1'172,698.8	1'029,533.8	87.8
1997		1'349,327.7	1'233,253.5	91.3
1998		1'510,300.0	1'492,545.7	98.8
1999		1'374,996.3	1'360,205.4	98.8
2000		2'096,873.9	1'568,145.4	74.7
2001		1'625,765.6	806,385.7*	99.2

* La ejecución corresponde al primer semestre.

Fuente: Fidelist, 2001.

Es importante estudiar la ejecución en consonancia con los dos destinos principales del presupuesto. Así pues, los resultados indican que para el año 2000 solo se ejecutó 72.4% de los recursos destinados para el subsidio al consumo (beneficiarios), al establecer la relación de lo asignado y lo ejecutado (1'294,288.5/1'785,704.5). A la par, se observa 88% para operación e inversión (273,856.5/311,169.4). Esto, a todas luces, reflejaba problemas de gestión y, por tanto, una primera razón para el incumplimiento de los objetivos propuestos. Esta situación

puede obedecer a muchas causas, como sobreestimación en el ejercicio presupuestario, ya sea por los movimientos en el padrón de beneficiarios o porque la frecuencia de retiro del subsidio presenta fluctuaciones significativas. De igual modo, el presupuesto asignado, en esencia, se destinaba a cubrir dos rubros: subsidio a la operación y la inversión, y subsidio al consumo. El primero permitía a las agencias estatales operar y administrar el programa en cada una de las 32 entidades federativas. El segundo correspondía a los pagos efectuados a los industriales de la tortilla, quienes proveían la tortilla en las diferentes localidades donde hacía presencia el programa.

8.3 INVERSIONES

Para la operación del programa se demandaron dispositivos de almacenamiento electrónico y sistemas automatizados, los cuales permitieron el control sobre el funcionamiento del programa. Para tal efecto, se entregó la tarjeta plástica con información del beneficiario y del proveedor de las tortillas. Desde luego que para la lectura de la tarjeta fue necesaria una terminal con tecnología de lectura de código de barras en la que se registraba la entrega diaria de cada kilogramo de tortilla. Por tanto, las inversiones del programa corresponden al costo de adquisición de dichas terminales. El costo histórico promedio fue de US\$310 por unidad (dato proporcionado por la oficina de planeación). El número mínimo de equipos lectores se encontró en Baja California Sur con 49 y, el máximo, en el Estado de México con 766.

Teniendo en cuenta que cada establecimiento poseía un equipo lector y que el número de industriales al 31 de junio de 2001 era de 9,038, el costo total de inversión ascendió a \$26'336,732.² Por último, el programa no requirió inversiones adicionales por concepto de construcción de obras de infraestructura para las sedes de las gerencias estatales, tampoco en lo concerniente a costos de dotación de mobiliario o equipos de oficina.

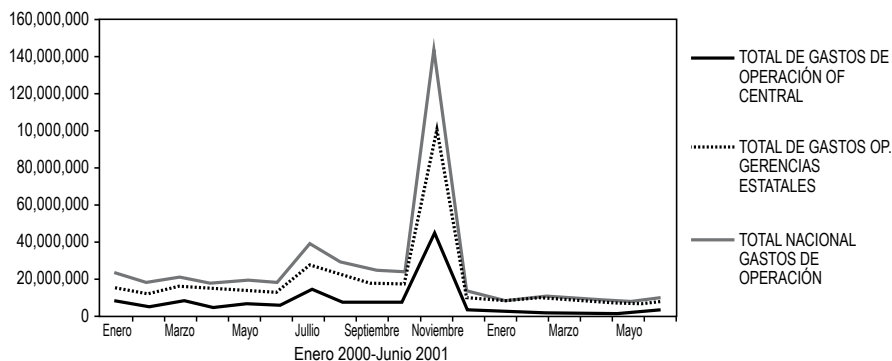
8.4 COSTOS DE OPERACIÓN

Los costos de operación estaban representados por todas aquellas erogaciones adelantadas por las actividades de administración, apoyo logístico, monitoreo y control, necesarias para el funcionamiento apropiado del programa. Para ese entonces, el Programa de Subsidio al Consumo de la Tortilla determinó una estructura organizacional de 32 gerencias estatales, haciendo presencia en todo el territorio nacional y una oficina central. Los principales costos de operación estuvieron clasificados en tres grandes rubros:

² La tasa de cambio observada el 3 de enero de 2000 fue de \$9.40 por dólar.

1. Servicios personales. Asociados con todas las partidas presupuestales destinadas para el pago de los salarios del personal adscrito a la institución y del personal eventual, y con todas las obligaciones derivadas de la contratación. Por tanto, se encuentran asignaciones para vacaciones, primas, gratificaciones, aportaciones al seguro social y, en general, todas las prestaciones sociales de ley.
2. Materiales y suministros. Consideran todos los gastos para insumos de aseo y limpieza, útiles de oficina y materiales complementarios para el mantenimiento de las instalaciones, más los gastos de combustible.
3. Servicios generales. Desembolsos ocurridos por concepto de servicios de agua, luz, postal, teléfono y arrendamientos de las oficinas desde las que se gestionaba el programa, así como el pago de algunos servicios externos de informática, estudios e investigaciones en beneficio del programa.

La estructura de los costos de operación indica que los servicios personales representaron en promedio 88%, seguido de servicios generales con 9% y 3% por concepto de materiales y suministros, respectivamente. Es importante observar el incremento sustancial de noviembre de 2000 por concepto del subsidio a la operación, ya que en la primera quincena se pagaron prestaciones de retiro, indemnizaciones y demás obligaciones, en virtud de la extinción del Fidelist (gráfica 8.2).



GRÁFICA 8.2. COSTOS DE OPERACIÓN

Fuente: elaboración propia con datos del Fidelist.

A la vez, es importante resaltar la participación del Estado de México en el subsidio a la operación que durante el año 2000 concentró más de 48% de los recursos destinados para tal fin, descendiendo su participación para el primer semestre del siguiente año al 30%. Sin considerar los costos de las oficinas centrales, las gerencias estatales de San Luis Potosí, Durango, Aguascalientes y Zacatecas, presentaron un mayor porcentaje de gasto de operación en

servicios personales en el mismo año, superior a 90%. A su vez, el Estado de México, Baja California Sur y Sonora gastaron un porcentaje superior a 12% en servicios generales, siendo las gerencias con los más altos gastos por este concepto.

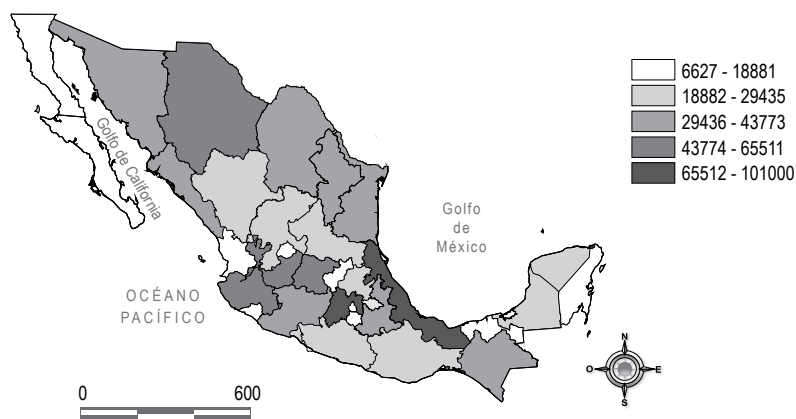
8.5 BENEFICIOS

Es importante tener presente que toda política pública, cualquiera que esta sea, genera beneficios no económicos. Los beneficios derivados del Programa de Subsidio al Consumo de la Tortilla se analizaron desde tres distintas posibilidades:

1. El número de familias atendidas mensualmente.
2. Los kilogramos de tortilla abastecidos cada mes.
3. El subsidio otorgado a las familias en pesos corrientes.

Conocer estos beneficios permite identificar las bondades del programa y sopesar su importancia.

8.5.1 Familias atendidas



MAPA 8.2. BENEFICIARIOS DEL PROGRAMA DE APOYO AL CONSUMO DE LA TORTILLA (NIVEL ESTATAL)

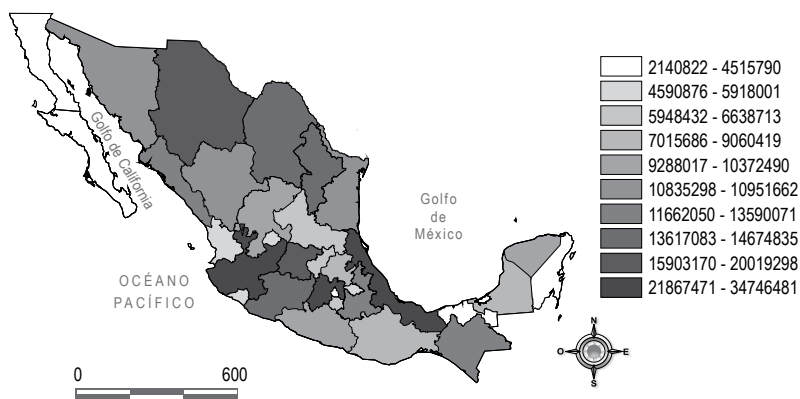
Fuente: elaboración propia con datos del Fidelist-Liconsa.

Una vez identificadas las colonias por parte del Fidelist, fueron seleccionadas las familias cuyas condiciones socioeconómicas estaban ligadas a la pobreza extrema. Este proceso se realizó evaluando las características de las viviendas, así como la posesión de enseres menores y el

perfil socioeconómico de los miembros del hogar. A partir de este ejercicio de focalización se estableció un padrón de beneficiarios en donde se atendieron un total de 1'139,781 familias en promedio a nivel nacional (de enero de 2000 a junio de 2001). Territorialmente, se observó que los beneficiarios tuvieron un mínimo de 7,069 familias en Baja California Sur y un máximo de 109,948 en el Estado de México. Considerando el total acumulado de beneficiarios durante el periodo de evaluación (20'453,061), la mayor proporción se encontró en la gerencia estatal del Estado de México, al representar 9.67%, seguido de Veracruz con 7.83% y Jalisco con 6.2% (mapa 8.2).

8.5.2 Kilogramos de tortilla abastecidos

Durante el año 2000, el programa suministró un total de 360'388,813 kilogramos de tortilla; de nueva cuenta, el Estado de México, Veracruz y Jalisco presentan la mayor concentración, seguidos de la Ciudad de México. Si bien, la provisión del subsidio era diaria, la verdadera regularidad del beneficiario para retirar el kilo de tortilla se puede establecer estadísticamente (mapa 8.3).



MAPA 8.3. KILOS ENTREGADOS. PROGRAMA DE APOYO AL CONSUMO DE LA TORTILLA (NIVEL ESTATAL)

Fuente: elaboración propia con datos del Fidelist-Liconsu.

8.5.3 Subsidio otorgado

Con relación directa a los kilogramos entregados, el monto del subsidio refleja, no solo, la cantidad transferida a nivel nacional, sino también, la diferencia del precio en cada entidad territorial. En el año 2000, el monto total del subsidio transferido alcanzó un monto de \$1.294'288.468,88 pesos, y para el primer semestre de 2001, el subsidio transferido llegó a \$751'100,855 pesos. El subsidio anual que cada familia beneficiada recibió fue de \$1.135,55 pesos.

8.6 INDICADORES DE EVALUACIÓN

Para los propósitos de la evaluación del programa se proponen como preguntas las siguientes: ¿cuánto cuesta atender una familia beneficiada?, ¿cuánto cuesta entregar un kilo de tortilla o una unidad de beneficio?, ¿cuánto cuesta transferir un peso de subsidio? A este tenor, la evaluación se centra en el diseño de indicadores desde la perspectiva del análisis costo-beneficio, garantizando así criterios objetivos para calificar el desempeño económico del programa. Con la información proporcionada por las directivas del programa sobre los costos de operación e inversión para cada una de las gerencias estatales y de los récords en kilogramos y pesos de subsidio del padrón de beneficiarios, se construyeron tales indicadores.

En primera instancia, la tasa de descuento equivale a la tasa de interés utilizada para actualizar el flujo de caja configurado para cada una de las gerencias estatales. Esta tasa refleja el valor del dinero en el tiempo y refiere sus usos alternativos. Para esta evaluación, la tasa de descuento considerada es la correspondiente al porcentaje de inflación acumulada durante los 18 meses del periodo de evaluación. La principal razón para adoptarla es por tratarse de un proyecto de inversión social implementado por agencias del Estado que, dada su naturaleza, no persiguen el lucro o un ejercicio de carácter financiero. La tasa de interés considerada debe proteger los recursos invertidos de procesos inflacionarios, manifestados por cambios en los precios relativos y con una consecuente pérdida del poder adquisitivo. A este tenor, las estadísticas oficiales del Banco de México indican que, de enero de 2000 a junio de 2001, se registró una inflación acumulada de 10.69%.³ Para efectos de la evaluación económica del programa, dicha tasa equivale a 0.85%, misma que se utiliza para descontar los costos mensuales del programa en cada gerencia estatal.

En segundo lugar, el Costo Mensual Equivalente (CME) servirá como referente para la construcción de los respectivos indicadores. Es conveniente recordar que para la obtención de este valor debe calcularse el Valor Presente de la inversión y los costos de operación para cada una de las entidades federativas bajo estudio, identificado como VPC. Posterior a ello, este valor se distribuye, de manera uniforme, en los 18 meses de evaluación, obteniendo así el Costo Mensual Equivalente que, en realidad, es una serie uniforme A , con $n = 18$ y constituye un costo por tratarse de una erogación o salida de dinero para atender las distintas obligaciones del programa. Matemáticamente, la relación entre el VPC y el CME puede expresarse como:

$$\text{CME} = (\text{Inv} + \text{VPC}) \left[\frac{i(1+i)^n - 1}{(1+i)^n - 1} \right]$$

³ Ver www.banxico.org.mx/cgi_inveco/sibamexbd.exe. La inflación es considerada como una variación porcentual respecto al mes anterior.

Lo anterior equivale a calcular una serie uniforme dado un valor presente. En la notación estándar de la ingeniería económica: A/P , con $n = 18$ y una tasa de interés donde n es el periodo de vida de la alternativa y la i es la tasa de descuento. Como se trata de construir indicadores de costo-eficiencia, dicho CME se relaciona con los respectivos beneficios obtenidos que, para este caso, se expresaron en familias atendidas, kilos de tortillas entregados o el subsidio otorgado expresado en unidades monetarias (pesos).

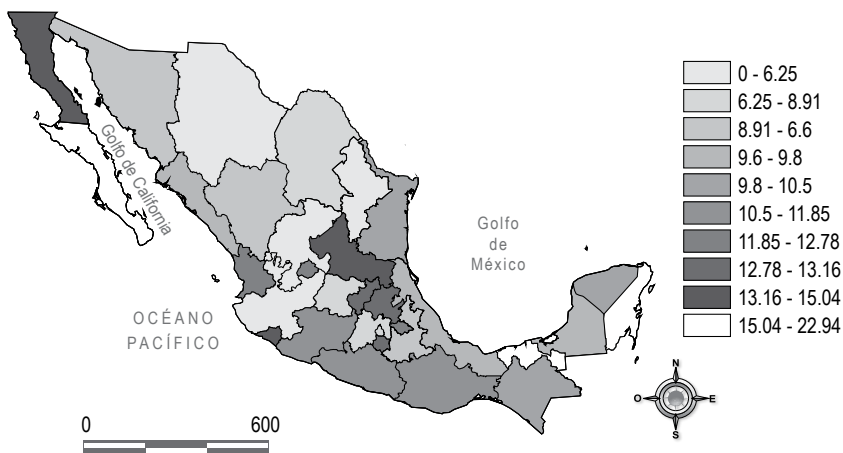
8.7 EVALUACIÓN COSTO-EFICIENCIA DEL PROGRAMA DE LA TORTILLA

8.7.1 Indicador de costo-eficiencia por familia atendida (C/E_i)

Este indicador resulta de dividir el costo mensual equivalente de cada gerencia o entidad federativa, entre el número de familias atendidas, en promedio, durante el periodo de evaluación:

$$\frac{C}{E_i} = \frac{\text{Costo Mensual Equivalente}}{\text{Familias atendidas mensualmente}} = \frac{CME}{F}$$

Los resultados oscilan entre \$5.64/por familia en Jalisco hasta \$22.92 en Baja California Sur. El promedio nacional de atención de una familia fue de \$11.21 por familia. El Estado de México –que concentra el mayor número de familias beneficiadas– presenta un desempeño de \$8.92 por familia, mientras que en Veracruz es de \$9.68. En el mapa 8.4 se representan los valores obtenidos para cada entidad territorial.



MAPA 8.4. INDICADOR COSTO-EFICIENCIA I. PESOS POR FAMILIA ATENDIDA (\$/FAMILIA)

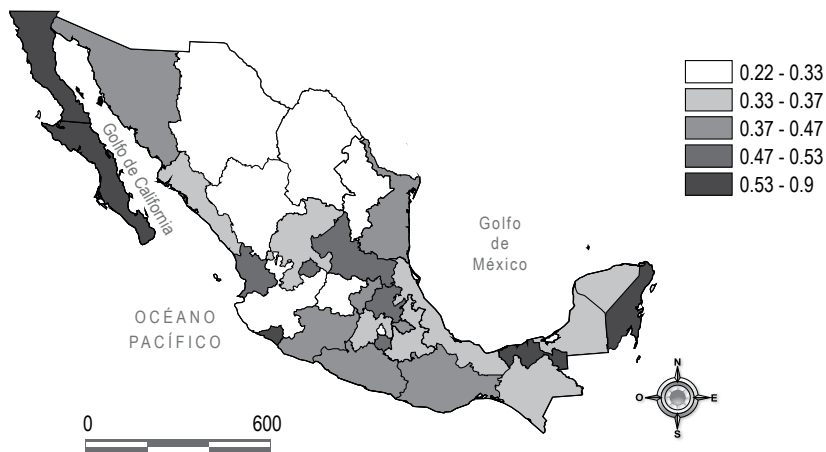
Fuente: elaboración propia con datos del Fidelist-Liconsu.

8.7.2 Indicador de costo-eficiencia por kilo entregado

Resulta de dividir el costo mensual equivalente de cada gerencia o entidad federativa entre los kilos, en promedio, entregados durante el periodo de evaluación.

$$\frac{C}{E_2} = \frac{\text{Costo Mensual Equivalente}}{\text{Cantidad de kilos entregados}} = \frac{CME}{K}$$

De nueva cuenta, Jalisco presenta el mejor indicador de costo-eficiencia por kilogramo (\$0.22), seguido de Nuevo León (\$0.26), Coahuila (\$0.29) y Chiapas (\$0.36) por kilo, cada uno. Las gerencias estatales que presentan los costos más altos son: Baja California Sur (\$0.92) y Baja California Norte (\$0.74). El precio promedio de entrega de un kilo de tortilla fue de \$0.43 por kilo. En el mapa 8.5 se representan los valores obtenidos por cada entidad territorial.



MAPA 8.5. INDICADOR COSTO-EFICIENCIA 2. PESOS POR KILO DE TORTILLA ENTREGADO (\$/KILO)

Fuente: elaboración propia con datos del Fidelist-Licons.

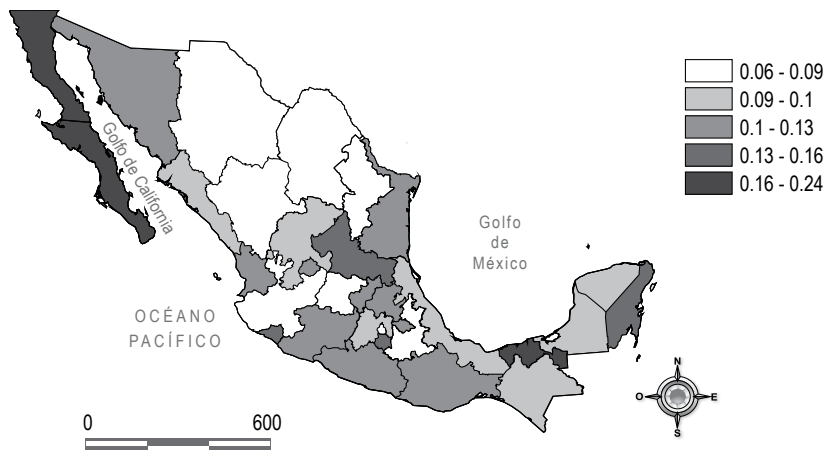
8.7.3 Indicador de costo-eficiencia por peso subsidiado

Resulta de dividir el costo mensual equivalente de cada gerencia entre el dinero entregado bajo la forma del subsidio durante el periodo de evaluación:

$$\frac{C}{E_3} = \frac{\text{Costo Mensual Equivalente}}{\text{Subsidio otorgado (\$)}} = \frac{CME}{S}$$

Los resultados indican que el rango de entregar un peso de subsidio oscila \$0.06 y \$0.24. Los niveles más bajos se encuentran en las gerencias estatales de Jalisco (\$0.06) y Nuevo

León (\$0.07), mientras que en los niveles superiores están Baja California Sur (\$0.24) y Baja California Norte (\$0.19). El promedio nacional para entregar un peso de subsidio fue de \$0.12 por peso. En el mapa 8.6 se representan los valores obtenidos por cada entidad territorial.



MAPA 8.6. COSTO-EFICIENCIA 3. PESOS POR PESO DE SUBSIDIO (\$/\$ SUBSIDIO)

Fuente: elaboración propia con datos del Fidelist-Liconsá.

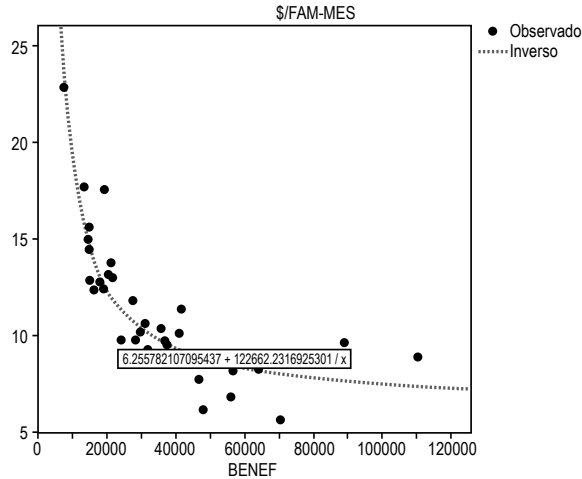
8.7.4 Relación entre los indicadores de costo-eficiencia

El diseño de los indicadores de costo-eficiencia permitió plantear nuevos interrogantes acerca de las relaciones que pudieran existir entre cada uno y el número de beneficiarios seleccionados en cada entidad. Para dicho modelo se asumió el número de beneficiarios como variable independiente y como dependiente cualquiera de los indicadores obtenidos. Esto significa que, para lograr cierto nivel de eficiencia, el programa debe tener un número adecuado de familias o beneficiarios que atender. A este tenor, se puede representar dicho modelo de la siguiente manera:

$$C/E_i = \beta_0 + \beta_1(\text{Número de beneficiarios}) + \varepsilon$$

Dado el comportamiento de los datos y el diagrama de dispersión entre el número de beneficiarios y los respectivos indicadores de costo-eficiencia, pueden observarse las relaciones inversas existentes en dicho comportamiento. Lo que indica que, a medida que crece el número de beneficiarios, el respectivo indicador disminuye (gráfica 8.3).

El coeficiente de correlación (relación entre variables) presenta los siguientes grados de relación: 0.673 entre beneficiarios e indicador de pesos por familia atendida, 0.646 entre pesos por kilo suministrado, y 0.644 entre pesos por cada peso de subsidio otorgado.



GRÁFICA 8.3. DISPERSIÓN DEL MODELO

Fuente: elaboración propia.

Es importante anotar que para hacer ajustes curvilíneos existen varias posibilidades. Asimismo, puede pensarse en modelos logarítmicos, de potencia, cuadráticos o de inversa. El modelo curvilíneo que mejor ajusta la curva es este último y se representaría de manera general como:

$$y = b_0 + (b_1/\text{beneficiarios})$$

Se puede afirmar que este modelo mejora el ajuste encontrado e interpreta mejor los datos observados. Algunas de sus características estadísticas se muestran a continuación:

Medida	Valor
Múltiple R	0.92790
R ²	0.086100
Adjusted R. Square	0.85636
Standard Error	1.31177

Para una mejor apreciación, la gráfica 8.3 presenta el modelo ajustado, anotando que inducir la eficiencia vía número de beneficiarios se logra, de forma rápida, dada la pendiente tan pronunciada a un nivel inferior de 20,000 familias. Esto refuerza la idea de estimar o determinar el número mínimo que una gerencia estatal debiera atender si se quiere mejorar el desempeño.

8.8 EVALUACIÓN GLOBAL DEL PROGRAMA

Cuando se integran los costos y los beneficios de las gerencias estatales con los costos de la oficina central, se puede evaluar el programa en su conjunto, calculando los respectivos indicadores. A este tenor se precisa:

- a) El valor actual neto, al ser una erogación, continúa siendo negativo y alcanza un monto de \$457'590,434.86.
- b) La relación costo–beneficio continúa siendo igual a cero.
- c) La TIR es de -70%.
- d) El CME es de -\$25'535,887.91.

Considerando este CME se construyen los indicadores de costo–eficiencia y globales, obteniendo:

$$C/E_1 = \$22.40/\text{familia atendida}$$

$$C/E_2 = \$0.87/\text{kilo suministrado}$$

$$C/E_3 = \$0.24/\text{peso subsidiado}$$

Se pone de manifiesto la significativa importancia de las oficinas centrales y el impacto económico que tienen para el programa en su conjunto. El costo promedio de atención a una familia se incrementa en más de 100%. Es importante recordar que el promedio nacional de atención de una familia fue de \$11.21 por familia. A la par, el precio promedio de la entrega de un kilo pasa de \$0.43 a \$0.87 pesos por kilo, lo que significa un incremento de 100%. Finalmente, las agencias estatales invertían en promedio \$0.12 para entregar un peso de subsidio, pero al incluir los costos de las oficinas centrales, este valor pasa a ser \$0.24/peso. Estos resultados reflejan la importancia de las oficinas centrales en la determinación de los indicadores de desempeño del programa como un todo y confirman la necesidad de acciones administrativas para atenuar dicha contribución.

8.9 SÍNTESIS CAPITULAR

Esta experiencia de evaluación económica sirve para reflexionar sobre las controversias que existen en torno a la eficiencia de los programas y políticas gubernamentales. A pesar de que los gobiernos no persiguen las utilidades económicas, es un imperativo para sus agencias hacer buen uso de los recursos del presupuesto público. La evaluación económica realizada

permitió establecer la utilidad de la técnica como también los niveles de eficiencia diferenciados territorialmente. La construcción de indicadores permitió resolver los interrogantes normales que pudiera enfrentar una autoridad: ¿cuánto cuesta atender una familia?, ¿cuánto cuesta entregar una unidad de beneficio?, ¿cuánto cuesta transferir una unidad monetaria de subsidio? Inquietudes imperativas para cualquier ordenador del gasto, así como para cualquier contribuyente, aunque no se planteen o se evadan sus respuestas. Una de los condicionantes importantes es la precisión y clasificación de la información financiera, económica y contable para emprender este tipo de ejercicios, algo que no es regular en el sector público, pero que para el caso de estudio resultó ser una de las grandes virtudes.

Una primera evaluación encontró diferencias entre el presupuesto programado y el ejecutado, que obedecen a las variaciones del retiro efectivo de la tortilla, por parte de los usuarios y denominadas fluctuaciones del padrón de beneficiarios. De igual manera, se reconocen grandes diferencias entre las gerencias estatales intervinientes, encontrándose algunas con niveles de eficiencia y efectividad notables. También es importante identificar los niveles mínimos de atención a usuarios en los que se alcanzan indicadores de desempeño relacionados con economías de escala para la provisión de los beneficios. Por último, es de destacar el impacto de los costos atribuidos a las oficinas centrales y el efecto de la liquidación organizacional en noviembre de 2000.⁴

REFERENCIA

Diario Oficial de la Federación. Reglas de Operación de los Programas del Sector de Desarrollo Social para el ejercicio del año 2000. https://dof.gob.mx/nota_detalle_popup.php?codigo=2052239

⁴ En el año 2000 y de acuerdo con el artículo transitorio del Decreto de Presupuesto de Egresos de la Federación, el programa de la Tortilla en su figura de Fideicomiso para la Liquidación del Subsidio a la Tortilla (Fidelist) se adscribe a Liconsa, S.A. de C.V. y se unifican las estructuras administrativas. En 2003 se reporta la disminución de las familias atendidas y desaparece en 2004, pues otros programas sociales lo sustituyen. <http://www.liconsa.gob.mx/wp-content/uploads/2015/07/MAN-ORG-GENERAL.pdf>

CAPÍTULO 9

Evaluación política

INTRODUCCIÓN

En su libro *Evidencia, argumentación y persuasión en la formulación de políticas*, Giandoménico Majone (1997) sentenciaba que, como muy bien lo sabían los políticos, la política pública estaba hecha de palabras. Esta premisa devela algo tan simple, pero que muchas veces no se admite, aunque sea evidente: toda política pública tiene motivaciones políticas. Esto es, promueven valores políticos y la aspiración implícita es lograr “poder” y “apoyo político”. También buscan adeptos, pretenden hegemonías y, definitivamente, van mucho más allá de las cuestiones técnicas y económicas. Asumirla como una acción neutra es profesar una ingenuidad infinita. Estas premisas permiten abrir un nuevo espacio para la evaluación en dimensiones no tradicionales, toda vez que implican una nueva concepción: las políticas como un discurso.

Las implicaciones de esta concepción son trascendentales para las discusiones ontológicas, epistemológicas y metodológicas. Asumir una política como un discurso abre nuevas posibilidades para la evaluación y exige nuevas consideraciones para juzgar sus asuntos. Las consecuencias también son importantes e imponen retos mayores para el evaluador, ya que le obligan a entender todo lo que entrañan las construcciones discursivas, sus valores, intenciones, ideologías, verdades y, muy seguro, una que otra mentira. El presente capítulo asume como tesis central que las cuestiones técnicas y económicas han opacado otras esferas de la función evaluativa, impidiendo que se exploren asuntos centrales y trascendentales de la acción y la confrontación política.

Si bien no existen muchas evaluaciones de esta naturaleza, se aspira que sea una de las futuras modalidades a considerar, pues los programas y las políticas tienen objetivos y pretensiones de tal naturaleza, que bien vale la pena estudiar. Inicialmente, se abordan las cuestiones propias del análisis del discurso.

9.1 LAS POLÍTICAS PÚBLICAS COMO DISCURSO

Una política puede asumirse como discurso, pues como se anticipó, cualquier programa o política es un evento o estructura comunicativa plena de valores, ideologías, teorías, buenas intenciones, en ocasiones explícitas y otras veces no. Sin embargo, hay intenciones no siempre develadas que deben ser estudiadas e interpretadas, pues motivan, condicionan y dan forma a las acciones y conductas sociales. Los discursos como las políticas, en todo caso, están soportados en complejos entramados de información que, usadas de manera estratégica, sirven a los fines gubernamentales y sociales para dirigir la acción colectiva y tramitar el conflicto social. Más allá de estas nobles intenciones, están el asunto de la hegemonía, los antagonismos, las supremacías y todos aquellos elementos del poder. ¿Cómo se diseñan las políticas?, ¿con qué intenciones lo hacen?, son interrogantes cruciales tanto para el analista como para el evaluador.

Conviene, también, reflexionar sobre quién tiene el control o la producción de tal discurso. Dicho de otra manera, estudiar quién produce el discurso, quién formula la política y cuáles son las características de la estructura discursiva. Desde la perspectiva decisionista, el discurso de la política es una competencia exclusiva de expertos y decisores, ubicados, por lo regular, en las élites políticas y difundidos de forma vertical. Es un discurso impuesto a, y asimilado y usado por la sociedad; en ocasiones, su verticalidad no permite la interlocución y, mucho menos, el juicio o la controversia. No está por demás recordar los postulados eastonianos cuando proponía que el estudio de la política, en singular, está interesado en comprender cómo las decisiones autorizadas son hechas y ejecutadas en una sociedad (Easton, 1957). Claro está, como bien lo planteaba el mismo autor, es un modelo muy “primitivo” y básico para entender la complejidad del mundo político con un marcado acento Estado-centrista, que se basa exclusivamente en el proceso decisorio. Por supuesto, diametralmente distinta es la perspectiva argumentativa, donde el discurso de la política puede, de alguna manera, ser el producto de una interacción –una deliberación–, con elementos de persuasión y retórica en espacios de debate y confrontación de los actores inmersos en dicho proceso. Desde la perspectiva argumentativa, la política pública es un discurso y, en último caso, podría considerarse una construcción social. Bajo estas consideraciones, el evaluador requiere conocer los fundamentos de análisis del discurso, pues tiene que incorporarlos como elementos evaluativos.

La verdad es que cuando se estudian las decisiones, ya sean estas individuales o colectivas, el tema de los valores siempre ha sido, es y será un foco de análisis. Ello, al parecer, deriva de los principios epistemológicos que pregonan los metadiscursos para generar conocimiento. El positivismo, por ejemplo, se declara una ciencia libre de valores, mientras que la ciencia social interpretativa y la crítica, ponen de relieve dichos valores y al propio servicio de la generación del conocimiento. Estas discusiones se han trasladado a las distintas disciplinas y las políticas públicas no son ajenas, por lo que también registran estas tensiones. La visión tec-

nocrática de las políticas acoge la objetividad como principio del consejo que el científico de las políticas otorga al gobernante. Dicha condición es cada día menos aceptada y resulta casi siempre inconveniente. Como lo sugiere Stewart (2009), se requieren nuevas normas en la ciencia de la hechura de las políticas, mismas que estén centradas en la claridad y en la explicitud acerca de las bases para los juicios que deben hacer los científicos, incluyendo los valores éticos y sociales de su rol. Para el autor, los valores en las políticas públicas son los fines valorativos incorporados e implementados por medio de las decisiones colectivas que se hacen mediante el proceso de la política pública. Más allá de ser el principio orientador de la acción colectiva, es el motivador como objeto mismo de la política; también el que da sentido de la realidad a los participantes como a los observadores y, en consecuencia, constituye un espacio emocional.

Postular una política pública libre de valores es negar su esencia y sus motivaciones. En consecuencia, se aconseja no purificar la ciencia y sí abandonar la pretensión del ideal de una ciencia libre de valores (Stewart, 2009). Una ciencia que se centre en la naturaleza del juicio, necesaria para que los decisores estén totalmente informados de la naturaleza científica del consejo recibido. La integridad de la ciencia no se asegura excluyendo los valores, sino asegurando que jueguen un papel aceptable en la ciencia. Sin importar su origen filosófico, jerárquico, espacial o social, toda política pública tiene una intencionalidad pues, al fin y al cabo, es un emprendimiento para atender algún asunto de la sociedad y buscar un mundo mejor. Con ello, se ratifica que las políticas no son neutras y mucho menos objetivas.

La objetividad, es decir, la centenaria pretensión positivista, vuelve a ponerse en duda o, por lo menos, a cuestionarse su logro. Desconocer los valores en las políticas públicas o en la acción política sería por demás ingenuo. Así, por ejemplo, el pensamiento clásico de la política siempre consideró su soporte en intereses, preferencias o ideologías, atributos estos basados axiológicamente. Tampoco los desarrollos teóricos contemporáneos han omitido esta consideración al asumir, de manera clara, la presencia de los valores en la acción política (Stewart, 2009; Meur, 2018 y Connor, 2019).

Una distinción sobre los valores es conveniente, ya que no pueden considerarse al mismo nivel ni pertenecientes al mismo dominio. Los valores que acompañan los metadisursos epistemológicos son distintos a los individuales, sus distinciones remiten a los debates de mitad del siglo XX. Los epistémicos incluyen la precisión predictiva, el poder explicativo, el espectro, la simplicidad (parsimonia), etc., bien diferenciados de los éticos como preocupación por la vida humana, reducción del sufrimiento, libertades políticas y normas morales (Stewart, 2009). Esta autora prefiere distinguir los valores en tres dimensiones y propone considerar una división en valores éticos, sociales y cognitivos, antes epistémicos.

Los valores éticos estarían enfocados en distinguir entre lo bueno o lo correcto, importantes cuando se trata de examinar las consecuencias del error o sopesar si los potenciales beneficios son más valiosos que los potenciales daños, o si algunos daños son irreparables o si

ostentan mayor peso (Stewart, 2009). Estos regularían el actuar científico en los procesos de generación de conocimiento (Douglas, 2009), tal es el caso de la experimentación sin consentimiento, el sufrimiento y la muerte de otros o la investigación en armas de destrucción masiva, etcétera.

Los valores sociales son muy cercanos a los éticos, pero que presentan una gran tensión por tratar temas como la justicia, la privacidad, la libertad, la estabilidad y la innovación. Los cognitivos suplantando los epistémicos y surgen como referente para asistir a los científicos en su actividad. La simplicidad, por ejemplo, sería un valor de una teoría, en tanto facilita trabajarla haciendo que el poder explicativo de un constructo teórico tenga mayor valor, en la medida que su aplicación lleva a mejores resultados. Otros valores cognitivos reconocidos son consistencia, espectro, precisión predictiva y productividad, comprometidos todos con las posibilidades del trabajo científico, animándolo e inspirándolo. En resumen, la generación de conocimiento también está ante la presencia de valores y, por tanto, el campo de las políticas públicas no escapa a su consideración.

Los valores son cruciales al determinar los propósitos perseguidos por la política pública. Tres razones validan esta afirmación:

1. Los sistemas de valores son cruciales para determinar si las políticas alcanzarán las metas enunciadas.
2. La legitimidad y su sistema de valores es en sí misma deseable para los resultados de la política.
3. En una estructura política, un rol importante de la política y las políticas consiste en tramitar las tensiones entre el sistema de valores.
4. Las decisiones de política y la conducta política de los actores afectan la cultura y los valores, retroalimentando y alterando el contexto de las futuras elecciones de política (Meur, 2018, p. 4).

La política se definía como “la distribución autorizada de valores en una sociedad” (Easton, 1965, p. 112). Sin embargo, la política es más que votos y lucha por el poder y, claramente, las políticas públicas se asimilan a las correas de transmisión de dichos valores. Toda política pública contiene valores, expresos o no, y desde la formulación del problema orientan su declaración. En la implementación dominan los valores de sus promotores o detractores. Una política pública puede definirse como una arena para el encuentro, colisión y prevalencia de valores; no reconocerlos encarna una miopía intelectual.

Los valores operan como filtros para que los individuos procesen información, minimicen la influencia externa y los guíen por medio de complejos entramados políticos. Los valores políticos son “concepciones generales abstractas acerca de lo deseable o lo indeseable,

del fin último de la vida humana” (Jacoby, 2006, citado en Connor, 2019, p. 3). Tienen una función evaluativa que permite a los individuos confrontar al mundo.

En torno a la visión instrumental de las políticas públicas y su problemática, se afirma que los valores son los mismos fines de las políticas, asimilables a la metas y obligaciones que se promueven como deseables (Rein, 2006). Se invocan metas como la seguridad, la equidad, la prosperidad, la libertad y el autogobierno, valores a la vez difíciles de conceptualizar, operacionalizar y alcanzar. Por el contrario, se convierten en motivo de controversia al abrir paso a diversas concepciones, medios para garantizarlos y maneras de evaluar su logro. También es necesario advertir que los valores, en ocasiones, se contraponen y entran en competencia o en conflicto. Se encuentran políticas públicas que buscan la equidad, pero que restringen la autonomía; inducen la participación, pero limitan la privacidad de algunos o fomentan la participación, pero reducen la racionalidad o el control tecnocrático. Lo cierto es que los valores están presentes y que, en ocasiones, superan con creces los beneficios materiales de las políticas y que su consecución no siempre se evalúa rigurosamente, por no tener conciencia de su existencia o por omisiones injustificadas.

9.2 EL DISCURSO COMO FOCO EVALUATIVO

El concepto de la política pública ha sido tan variado y difuso, llegando, en ocasiones, a ser contradictorio. Asumir una política como discurso, posibilita una nueva perspectiva epistemológica, distinta a la visión tecnocrática y Estado-céntrica que ha dominado la teoría de las políticas públicas. Implica también el uso de nuevas herramientas para el análisis, la investigación y, por supuesto, la evaluación. Subyace reconocer que la perspectiva tecnocrática tiene sus límites y que, si se aspira a mayor efectividad de la acción pública, habrá que añadir nuevas visiones. Sin ánimo de atender todas las cuestiones referidas al discurso, por resultar estas vastas, no unitarias y multidisciplinarias. A continuación, se presenta una breve referencia teórica a los asuntos discursivos identificando algunos elementos evaluativos.

La RAE asocia el discurso a una facultad racional con que se infieren unas cosas de otras, también a una construcción realizada, así como a una serie de palabras y frases empleadas para manifestar lo que se piensa o se siente. Adiciona otros criterios que dan cuenta de un razonamiento y conocimiento amplio acerca de una materia que se comunica o comparte con otros. Estas construcciones de ideas o conocimientos pueden constituir estructuras más elaboradas o abstractas para llegar a considerarse doctrinas, ideologías o tesis.

La teoría del discurso constituye un enfoque auxiliar para comprender cualquier política pública, centrado en el papel que representan las prácticas e ideas sociales significativas en la vida política (Howarth, 1995). Añade el autor, que su análisis pretende establecer “de qué manera los sistemas de significado [o ‘discursos’] configuran la comprensión que las personas tienen de sus propios roles sociales y cómo estos influyen en sus actividades políticas”

(Howarth, 1995, p. 125). Esta conexión con los significados establece una clara conexión con la ciencia social interpretativa y con sus modalidades como la hermenéutica, la fenomenología, el estructuralismo y la deconstrucción. Es un concepto amplio, ya que el discurso incluye todo tipo de prácticas sociales y políticas, así como las instituciones y organizaciones, y su intención es examinar de qué modo tales estructuras de significado hacen posibles ciertas formas de conducta, individual o colectivas. Para el analista, el desafío encarna comprender cómo se construyen los discursos que estructuran las actividades de los agentes sociales, cómo funcionan y cómo cambian a través del tiempo. Un desafío nada simple en su dimensión analítica.

Es menester recordar que el discurso como unidad de análisis ha transitado por distintas disciplinas. Así, por ejemplo, la lingüística, en su enfoque analítico, se limita a los pequeños segmentos del habla o de la escritura para determinar cómo empieza y cómo termina; también intenta estudiar de qué manera se introducen los temas, cómo se produce la narración y cómo se alternan las personas al conversar o intervenir. Un discurso (asúmase el lenguaje) puede entenderse como “un sistema de diferencias formales en el que la identidad de las palabras es puramente relacionada” (de Saussure, 1916, en Howarth, 1995, p. 129). Tal autor divide las unidades lingüísticas entre “significantes” y “significados” y, desde esta perspectiva, el discurso debe entenderse como una cuestión meramente textual. Más allá de esta perspectiva, el discurso es un evento cognitivo y social que se construye para representar no solo el texto sino también el contexto, resaltándose la relación pragmática que se establece entre estos dos elementos (van Dijk, 1977). Los estudios del discurso se volvieron relevantes en la década de los años setenta desde la sociolingüística, la etnometodología, la microsociología y la psicología, con variantes e interacciones entre estas disciplinas (van Dijk, 1998). El discurso tiene una función primordial en la expresión discursiva y, más importante, en la reproducción de las ideologías.

En los estudios políticos, la teoría del discurso es la que permite la comprensión de la vida política, es decir, del fenómeno y la conducta política, interesados en particular, en “el papel de las prácticas e ideas sociales significativas en la vida política” (Howarth, 1995, p. 131). Desde la dimensión política existe una clara intención por entender conceptos políticos como antagonismo, actuación, poder y hegemonía, pues, al fin y al cabo, “los discursos adquieren su identidad mediante el trazado de fronteras políticas y la construcción de antagonismos entre amigos y enemigos” (Howarth, 1995, p. 131).

Se reconoce también un vínculo fuerte entre el análisis del discurso y el posmodernismo, con notables exponentes como Michel Foucault, Ernesto Laclau, Chantal Mouffe y Jacques Derrida, entre otros más. En sus desarrollos, destacan la importancia de las palabras en la historia, el contexto, las alocuciones, las instituciones, las prácticas y las identidades. “Los discursos son históricamente contingentes y se construyen políticamente”; asimismo, afirman la supremacía de las prácticas políticas, en particular, en la configuración de las identidades (Laclau y Mouffe, 1990 en Howarth y Stavrakakis, 2000, p. 5). Estos autores sustentan su

análisis en la concepción de lo social como un espacio discursivo, mismo que posibilita las relaciones de representación. También reconocen, en el pensamiento contemporáneo, tres corrientes de pensamiento notables del siglo xx: fenomenología, filosofía analítica y estructuralismo.

9.3 LAS DIMENSIONES POLÍTICAS DE LA EVALUACIÓN

Desde su doctrina teórica y factual, la ciencia de las políticas suponía un fuerte compromiso con la democracia y con la realización de la dignidad humana, según los postulados lasswellianos de mitad del siglo xx. A pesar de ello, el seguimiento de las acciones gubernamentales, en algunos contextos, revela que estos preceptos no siempre son considerados como prioritarios. En ocasiones, la democracia y la dignidad humana están supeditadas a otros atributos como la eficiencia o el apoyo electoral. Debido a ello, tampoco las prácticas evaluativas de las políticas se consideran, existiendo así un déficit evaluativo y afectando las relaciones de los ciudadanos con autoridades y sistemas de gobierno. Con dificultad se encuentran investigaciones evaluativas que versen sobre los valores políticos, la rendición de cuentas, la responsabilidad, la transparencia, los *stakeholders* o la participación. Y es que los principios y valores del buen gobierno, por lo regular, no se incluyen como ejes centrales, pero tampoco son muy populares las estrategias metodológicas para emprender evaluaciones de esta naturaleza. Así las cosas, puede afirmarse que la evaluación no ha estado a la altura del compromiso de la dimensión política de las políticas.

Los analistas de políticas deben estar más atentos a los efectos ulteriores de las acciones gubernamentales; en especial, aquellos que promueven desequilibrios en los poderes y todas aquellas situaciones degenerativas que pudieran afectar o dañar la democracia (Ingram y Schneider, 2007). Dado que el análisis de las políticas es un ejercicio, en esencia normativo, los valores de la democracia debieran tener una especial consideración y atención en dichas actividades, según lo afirmado por las mismas autoras. Sin duda, ello desafía el tradicional enfoque de las políticas y, por supuesto, la tarea de los analistas.

Más allá de las cuestiones técnicas planteadas, “la evaluación posee atributos que deben ser reconocidos adicionalmente y valorados desde una dimensión política” (Carrillo, 2009, p. 59). En este sentido, evaluar es un acto político con consecuencias políticas. Las esferas políticas de la evaluación no han sido del todo reconocidas por sus practicantes o promotores. Por tal motivo, se presentarán algunas ideas sobre:

[...] los elementos políticos que dan contorno a la evaluación, advirtiendo que por su naturaleza compleja y extensa merecerían un mayor espacio para su tratamiento. Se reconocen de manera central el diálogo, la participación, la rendición de cuentas, la transparencia, como elementos centrales de la buena gobernación (*good governance*) (Carrillo, 2009, p. 59).

Dicho sea de paso, no constituyen novedad alguna en los estudios políticos por ser impulsados desde las agencias multilaterales.

9.3.1 Democracia y políticas públicas

La ciencia de las políticas de la democracia fue el término acuñado por Lasswell (2000) en su empeño por (re)crear esta disciplina. Para él, las políticas públicas deberían propiciar una relación simbiótica con los gobiernos democráticos, a fin de resolver los problemas de vivir en sociedad. Sin embargo, como se sabe, el autor hacía “referencia a su democracia”, la representativa, referente a todas luces válido, pero insuficiente. A decir verdad, las políticas públicas pueden y deben rebasar la visión Estado-centrista, como lo vienen promulgando algunos tanques de pensamiento y como lo exigen los actores sociales, pues los gobernantes pueden ser la fuente de la exclusión y la arbitrariedad (incluso quienes se autoproclaman demócratas).

Si las políticas públicas tratan los problemas de vivir en sociedad en un sistema democrático, resulta pertinente evaluar el desempeño gubernamental. Esto es, determinar la efectividad y el impacto de sus acciones sobre: el sistema político, los valores democráticos y las relaciones con los demás actores. Por ello, como se verá más adelante, es necesario estudiar la calidad de la democracia, comparar sus alcances y proponer una taxonomía a la luz del pasado de sus sistemas políticos y de su estado actual. Proponer indicadores o contribuciones de las políticas públicas para los atributos de la democracia complementaria, por tanto, esa necesidad de monitorear los avances hacia un mundo mejor.

9.3.2 Deliberación y evaluación

La evaluación, en su esencia, es un proceso con gran vocación por la deliberación y el diálogo (consustanciales). Privilegia “la comunicación, fomenta el conocimiento y la comprensión detallada de la acción pública ejercida en torno al tratamiento de los problemas de vivir en sociedad” (Carrillo, 2009, p. 59). Por tanto, cualquier ejercicio evaluativo concitaría el encuentro y el intercambio entre las partes. “El diálogo, por tanto, se constituye en ese mecanismo o medio para canalizar aquella información generada de un acto evaluativo y debe ser difundida, si se quiere pensar en la utilidad de la función” (Carrillo, 2009, p. 59). “[L]a evaluación encarna un proceso comunicativo que utiliza el diálogo para reflexionar sobre aquellos hallazgos del ejercicio” (Carrillo, 2009, p. 59). No es posible evaluar sin dialogar y sin comunicar, ya que media un acto político, de confrontación de ideas y valores, de intercambio y de negociación.

Para la evaluación de las políticas, el diálogo y la deliberación se consideran, no solo, “un deber ético-político del conjunto de ciudadanos que integran la comunidad política” (Baca, 1996, en Carrillo, 2009, p. 60), y el medio para mejorar el desempeño de las agencias en-

cargadas de administrar los distintos programas cobijados por la acción gubernamental (Carrillo, 2009). Para los asuntos de la democracia y de la acción colectiva, el diálogo y la deliberación son principios para la atención de los asuntos públicos, rechazando así la idea que el análisis de política pueda ser una directa aplicación de las técnicas científicas. Desde esta posición, se toma distancia del decisionismo, aquel enfoque que privilegia la toma de decisiones por aquellos que detentan el poder, que creen tener la verdad y que monopolizan el proceso decisional de las políticas.

La democracia deliberativa es, por tanto, una modalidad que privilegia estos elementos para abordar el proceso de las políticas públicas y, en consecuencia, la evaluación. Es una perspectiva incluyente, que exige la creación de espacios para el debate público, la participación y la construcción de las políticas. Denominado como el giro argumentativo, esta perspectiva configura una nueva orientación para el análisis de las políticas y de la planeación (Fisher y Gotweiss, 2012). Un cambio más allá del dominante empírico y analítico enfoque a la solución de problemas hacia uno que estudie el lenguaje y la argumentación como dimensiones esenciales de la teoría y el análisis en la hechura de las políticas públicas y la planeación. Basado en la teoría crítica de Jürgen Habermas, con su crítica de la tecnocracia y el cientificismo, y su trabajo sobre la acción comunicativa, el giro argumentativo ofrece un enfoque alternativo al estudio de las políticas.

Desde 1993, esta perspectiva analítica se ha venido expandiendo para incluir trabajos sobre análisis del discurso, deliberación, democracia deliberativa, jurados ciudadanos, *governance*, expertos, investigación participativa, conocimiento tácito y local, planeación colaborativa, usos y roles de los medios de comunicación y métodos interpretativos, entre otros. Reconociendo que el proceso de las políticas está constituido y mediado por las prácticas comunicativas, el giro argumentativo intenta entender los procesos de hacer política y las actividades analíticas de la investigación en sus propios términos. El enfoque trabaja para entender lo que hace el analista de políticas, las formas de comunicar sus hallazgos y consejos, y su uso por parte de los receptores. Desafía la creencia de que el análisis de políticas está libre de valores y toma la argumentación como su punto de partida. Busca entender las relaciones entre lo empírico y lo normativo, en la medida que configuran el proceso de la argumentación de las políticas.

Asumiendo la propuesta del giro argumentativo como una realidad, la evaluación debe considerar las formas para juzgar lo obtenido a través de los espacios de discusión. En consecuencia, debe establecer los indicadores requeridos para valorar las ganancias de los procesos de encuentro, diálogo, deliberación y debate, pues resultan definitivos para la armonía, convivencia y consolidación democrática. Algunos indicadores posibles podrían versar sobre: espacios de debate propiciados por las políticas públicas, argumentos y evidencias presentadas y todas aquellas instituciones establecidas para la toma de decisiones inclusivas.

9.3.3 Participación y evaluación

“Según sea la modalidad que se utilice para emprender una evaluación, la participación puede presentarse con un desarrollo incremental sustantivo. Es bien sabido que algunas estrategias consideran el involucramiento de todas las partes inmersas en la política pública” (Carrillo, 2009, p. 60). Algunas modalidades de evaluación promueven la participación de todos los actores e invitan a sumar perspectivas diversas, pero vale la pena recordar el significado de “participar”. Según la definición del diccionario de la lengua española, da cuenta de lo que se dice de una persona: tomar parte en algo, referido a compartir, pareceres o comunicar o dar parte de algo. Invoca también un asunto de pertenencia, al hacer parte de una organización a un individuo. “Así, entendida la participación, puede ser un insumo necesario para evaluar los programas gubernamentales, toda vez que recoge las visiones de sus actores y los vuelve sujetos de la política” (Carrillo, 2009, p. 60).

Desde 2003, el Congreso de las autoridades locales y regionales de Europa dictó una carta acerca de la participación de los jóvenes a nivel local y regional. Desde entonces, se reconoce un marco que caracteriza las condiciones necesarias para que tomen lugar y están recogidas en el acrónimo RMSOS (por sus siglas en inglés). Inicialmente, es necesario tener el derecho (R), algo obvio, pero no extraño de reconocer; se requieren medios (M), espacios (S), oportunidades (O) y, cuando sea necesario, el apoyo (S) para participar e influir en las decisiones (Grace y Grace, 2017). Participar es el proceso de compartir decisiones que afectan la propia vida y la de la comunidad que se habita, es un medio para la construcción de la democracia y un estándar o indicador para medirla. Siendo así, es posible emprender acciones y actividades para contribuir en la construcción de una mejor sociedad.

Estas nuevas condiciones tienen implicaciones para la gestión de las políticas, pues en la actualidad, el ciudadano no está dispuesto a seguir siendo objeto de la acción pública, ni a continuar desempeñando un papel pasivo de simple receptor de acciones, exige que se le reconozca como sujeto de la acción pública y que su papel sea dinámico y activo. En términos simples, desea tener mayor participación en las actividades y decisiones de los gobiernos; exige la calidad, oportunidad y atención que encuentra en el sector privado; y cada vez quiere sacar un mayor provecho del dinero que aporta bajo la figura de impuestos, tasas, contribuciones o tarifas (Carrillo y Toca, 2007, en Carrillo, 2009, p. 61). Abandona el lamentable rol de súbdito y emerge con una vocación participativa, reclama un reconocimiento en la relación con sus autoridades y demandando mejores niveles de desempeño (Carrillo, 2009, p. 61).

La insatisfacción ciudadana es, por tanto, el foco de atención académica y de preocupación gubernamental por ocasionar e incrementar la turbulencia social, otorgando mayor complejidad e incertidumbre en la acción colectiva y traumatizado, en buena forma, el avance de algunas sociedades y grupos sociales. Interrogantes del tipo, ¿cuántas personas se involucraron?, ¿cuántas veedurías se conformaron? ¿cuál(es) fue(ron) el(los) rol(es) de los involucrados?,

llevan al establecimiento de algunos indicadores que revelan la participación en la evaluación y en el control de las acciones públicas.

9.3.4 Transparencia y evaluación

En la actualidad, lo público debe gozar de visibilidad y publicitación, procesos crecientes y propios de las sociedades más avanzadas. “En efecto, las nuevas actitudes ciudadanas, los nuevos roles emprendidos, así como los fenómenos de internacionalización y globalización han impulsado mayores aperturas y compromisos con el arbitrio de los asuntos públicos” (Carrillo, 2009, p. 61). Existe cierta supremacía de lo público sobre lo privado “que significa el aumento de la intervención estatal en la regulación coactiva del comportamiento de los individuos y de los grupos infraestatales, o sea, el camino de la emancipación de la sociedad civil respecto al Estado [...]” (Bobbio, 1997, en Carrillo, 2009, p. 61).

Estas nuevas tendencias han impactado los modos o estilos de hacer políticas públicas, de manera general, y de evaluar de manera particular. Las relaciones entre gobierno y sociedad se han horizontalizado y en algunos casos se han invertido, dejando de lado los modelos tecnocráticos y verticales para manejar la acción pública (Carrillo, 2009, p. 61).

Las tendencias mencionadas hacen de la apertura un efecto inevitable y propicio para que los decisores privilegien el interés público, trayendo consigo novedosas responsabilidades para las autoridades gubernamentales. La autoridad central ya no decide discrecionalmente y, en su lugar, se convierte en árbitro y promotor de la solución de conflictos del vivir en sociedad. El gobernante ya no es el protagonista principal, sino un agente más. Fue la legal-burocracia la que incapacitó al gobierno para desempeñar su papel público, de modo que el ciudadano tomó distancia, afectándose con ello la imagen y la legitimidad de las burocracias. En particular, el ciudadano latinoamericano tomó conciencia de los vicios y delitos administrativos (tramitomanía, corrupción, despilfarro, etc.), desencantándose de sus representantes. El ciudadano, de forma colectiva, empieza a exigir del sector público mayor receptividad, sensibilidad, reestructuración y adaptación de los organismos gubernamentales frente a las cambiantes necesidades ciudadanas. Reclamando en paralelo un estatus digno y privilegiado y ser los protagonistas de la actuación pública (Carrillo y Toca, 2008). “Al parecer, individuos y grupos sociales rechazan cada día más las relaciones verticales, y quieren superar las denominaciones de ciudadano-súbdito o, como ciudadano-electoral, recuperar la esencia de su condición ciudadana” (Carrillo y Toca, 2008, p. 6). Por tanto, se resignifican y reconfiguran los vínculos ciudadano-gubernamentales, en la medida en que los protagonistas reconocen que la ciudadanía ocupa un estatus mayor que el otorgado por la constitución para trascender a las dimensiones social y cultural, responsables de la identidad y pertenencia ciudadana (Kymlicka y Norman, 1994, en Carrillo y Toca, 2008).

Es la organización de minorías en colectivos, la que ha llevado a reconceptualizar los componentes de la ciudadanía. El ciudadano demanda ser protagonista de la escena pública y gozar de mejores condiciones de vida (Carrillo y Toca, 2008). El control y la denuncia social garantizan la transparencia, de modo que los evaluadores de la dimensión política tendrían que considerar las denuncias, procesos de investigación y recomendaciones para mejorar los procedimientos y garantizar la pulcritud en el uso de los recursos públicos.

9.3.5 Rendición de cuentas y evaluación

El también denominado *accountability* es un requisito para aquellos gobernantes que buscan mejorar su actuación y, de paso, tener relaciones armoniosas con su contraparte ciudadana; ello, en la medida que quisieran permanecer en el cargo. Reconocido también como principio democrático que pone límite a los poderosos, exigiéndoles ser responsables de decisiones y acciones. La rendición de cuentas aplica a distintos planos como el político, el administrativo, el moral y el legal. Sin duda, la evaluación puede aportar al cumplimiento de este principio democrático (Carrillo, 2009).

La vigilancia, el control y el refrendo o la sustitución de representantes estatales, por parte de los ciudadanos-electores, es propia de los gobiernos democráticos. Esta fórmula es la que otorga el poder ciudadano y privilegia la solución de conflictos y, con ello, la armonía y la convivencia sociales. Es una tendencia que perdurará en el tiempo, especialmente, por la decepción y frustración con algunas democracias. La propagación de la popularidad de la rendición de cuentas constituye una oportunidad para las sociedades que viven bajo regímenes opresores (Carrillo, 2009).

[L]a democracia actual, como la evaluación de la acción pública, empodera a los ciudadanos, ayuda a superar la exclusión y contribuye al buen gobierno, y puede convertirse en herramienta que aporta al manejo responsable de los recursos públicos y al destierro de los abusos del poder (Carrillo, 2009, p. 63).

Las veedurías populares y demás mecanismos de control social están diseñados para que las autoridades y los responsables del manejo de recursos públicos, presenten informes de gestión en espacios sociales. Indagar sobre interrogantes del tipo ¿contempla el diseño de políticas públicas espacios de control?, ¿garantiza los mecanismos necesarios para vigilar y rendir cuentas?, nos permite avizorar instrumentos y elementos para una evaluación de carácter político.

9.3.6 Buena gobernanza y evaluación

El *good governance* o buena gobernanza tiene relación con la demanda de mayor eficacia del Estado, ya que intenta responder las demandas de las sociedades altamente complejas, infor-

madas, activas, hastiadas de organismos que entregan menos calidad, pero que gastan más recursos y frustradas con las burocracias descontroladas y cooptadas por individuos insensibles e incapaces (Carrillo, 2009).

Los gobiernos deben abandonar relaciones jerárquicas de cara a la sociedad y la iniciativa privada, siendo dos las opciones: compartir la autoridad con otros generando entornos plurales y permitir que el mercado domine sin reclamo alguno. “La diferencia clave entre estos dos escenarios es la clase de autoridad legitimada para regular la sociedad, así que el buen gobierno puede ser construido sobre uno de los dos escenarios” (Carrillo, 2009, p. 63).

La eficacia, la promoción y la regulación derivada del Estado constituyen verdaderos desafíos en un intento por repensar su papel. Redefinir la responsabilidad pública de sus agencias es la primera acción, a fin de incorporar la participación de la comunidad en la prestación de bienes y servicios comunes. Ello ampliaría las capacidades de promoción y emprendimiento eficientes de acciones de interés común. Entender la buena gobernanza se logra reconociendo los rasgos de la mala gobernanza (Carrillo, 2009). Según el Banco Mundial, el mal gobierno se reconoce por:

- 1) un fracaso en hacer una clara separación entre lo público y lo privado, lo que deriva en una tendencia a orientar recursos públicos para el beneficio privado; 2) para establecer un marco legal predecible que promueva el desarrollo, o la aplicación arbitraria de las leyes y reglas; 3) establecimiento de prioridades inconsistentes con el desarrollo o que promuevan la mala distribución de los recursos; 4) presencia excesiva de reglas, regulaciones, etc., las cuales impiden el funcionamiento de los mercados y, finalmente, 5) un proceso de toma de decisiones poco transparente (Carrillo, 2009, pp. 63-64).

En consecuencia, los atributos de la buena gobernanza perfilan la acción gubernamental transparente, responsable, equitativa, capaz de dar respuesta a las necesidades sociales, eficiente y eficaz. Su logro dependerá de la comunidad de su interés por los asuntos gubernamentales, la veeduría de las decisiones y del ejercicio del gobierno. En otras palabras, de su activa participación y del control sobre el proceso de políticas aportarán a su inclusión política en los asuntos de gobierno (Carrillo, 2009).

9.4 LA DEMOCRACIA Y SUS INDICADORES

La democracia ha sido considerada uno de los valores universales acentuado en el siglo xx (Sen, 1999), corroborando así que el espacio democrático es propicio para la evaluación de las políticas públicas. Dado que la democracia se populariza como forma de gobierno, las investigaciones han estado atentas a la evolución de sus distintos aspectos, atributos, modalidades, etc., mismos que, en resumidas cuentas, pueden agruparse en los asuntos de la cali-

dad. La medición de algunos fenómenos sociales, aunque desafiante, constituye un pilar, no solo para la investigación sino una estrategia para mejorar el desempeño. Por ello, la revisión de algunas iniciativas, ayuda al reconocimiento de algunos atributos, indicadores y técnicas para su evaluación.

Un sistema de indicadores conviene a cualquier evaluador al brindar, a las agencias y a los *stakeholders*, herramientas para evaluar la acción política y, de alguna manera, ganar adeptos para que la evaluación sea un imperativo de la acción pública. No es una cuestión simple medir estos atributos, pues constituye un reto metodológico; sin embargo, aun así, en la historia pueden rastrearse muchas iniciativas de investigación política.

Como se puede constatar, los índices están de moda. Las agencias multilaterales, los centros de investigación y algunas instituciones educativas están contribuyendo a la tarea de estudiar la evolución de asuntos cruciales como el bienestar, el desarrollo, la transparencia, la democracia, la paz, etc. Se cuenta así con opciones interesantes para estudiar estos parámetros, conocer sus avances y retrocesos, y hacer investigación comparativa. No resulta muy claro el uso dado, sobre todo, por parte de las autoridades gubernamentales, pero aun así, pueden despertar la curiosidad por la medición de fenómenos y condiciones abstractas de la vida social. Asumiendo que la democracia es el derrotero político para perseguir por las políticas públicas. A continuación, se reseñan algunas iniciativas para constatar su estado y evolución.

9.4.1 La medición de la democracia

Si la teoría sobre la democracia impone un reto, su medición es un desafío mayor. Giovanni Sartori (1993) uno de los destacados estudiosos de la democracia, encuentra justamente dificultades a la hora de discutir sobre esta aspirada y escurridiza forma de gobierno. Plantea que la teorización, la práctica y su alcance son difíciles, sumado a que el camino democrático se encuentra lleno de trampas conceptuales, etimológicas y empíricas. En efecto, los estudios de la democracia presentan una primera dificultad, en tanto dicha forma de gobierno tiene una doble concepción: la positiva y la normativa. Más allá de lo que es y lo que debería ser, no sobra remarcar que sus modalidades son variadas, otorgando mayor complejidad a su estudio e indagación, encontrando la democracia electoral, participativa, deliberativa y radical a cada modalidad con sus propios defensores. El espacio donde ocurre también abre paso a la clasificación, por ello es común escuchar de democracia política, social y económica. Por si fuera poco, los niveles también dan lugar a la democracia internacional, nacional y local.

Evaluar y medir el progreso de la democracia nacional se reconoce como un esfuerzo académico y profesional para contar con guarismos que den cuenta de su comportamiento a lo largo del tiempo. Los métodos han sido diversos, desde escalas y categoría sobre las prácticas democráticas hasta percepciones basadas en encuestas de opinión pública (Beetham *et al.*, 2009). Ha prevalecido un afán por la comparación y la evolución, aspirando a que las socie-

dades y sus representantes actúen, en consecuencia, corrigiendo el rumbo de los acontecimientos. A riesgo de parecer muy básico e injusto, se presentan las metodologías más populares: Latinobarómetro, *Democracy Index (The Economist Intelligence Unit)* y *Global State of the Democracy* (SoD, IDEA). Estas proporcionan algunas dimensiones e indicadores para el estudio de la democracia en sus distintas dimensiones y ofrecen algunas ideas para que los evaluadores de políticas identifiquen ámbitos de interés político y construyan los respectivos indicadores.

Latinobarómetro “investiga el desarrollo de la democracia, la economía y la sociedad en su conjunto, usando indicadores de opinión pública que miden actitudes, valores y comportamientos” (Cepal, 2015). Son cerca de 20,000 encuestas aplicadas en 17 países de Latinoamérica que miden grados de satisfacción y de apoyo. Sus resultados se presentan como porcentajes de la población; por ejemplo, 49% de los encuestados en 2021 prefiere la democracia sobre otra forma de gobierno y 62% rechaza un gobierno militar (<https://www.latinobarometro.org>). Como se observa, es una metodología muy básica para establecer los porcentajes poblacionales en favor o en contra de algunos atributos de su gobierno, economía y sociedad.

Por otra parte, la Unidad de Inteligencia de *El Economist* propone un índice para lograr una imagen instantánea de la democracia mundial, considerando 165 estados independientes y dos territorios. La construcción de una medida de la democracia está basada en cinco categorías, a saber: proceso electoral y pluralismo, libertades civiles, funcionamiento del gobierno, participación y cultura política. Estas categorías recogen 60 reactivos de distinta naturaleza, que puntúan entre 1 si hay presencia, 0.5 si existen zonas grises o críticas y 0 si está ausente. En ocasiones, considera los indicadores de la encuesta mundial de valores (*World Value Survey*) para obtener un índice global en una escala de 0 a 10, derivado de un promedio simple de las cinco categorías. Este índice es usado para ubicar un país dentro de uno de los cuatro tipos de regímenes de democracia: plena, si su puntaje está entre ocho y diez; defectuosa, entre seis y ocho; híbrida, entre cuatro y seis, y régimen autoritario, menor a cuatro.

Otro destacado estudio es el del Instituto Internacional para la Democracia y la Asistencia Electoral (IDEA, por sus siglas en inglés). Esta iniciativa sueca es una de las fuentes más reconocidas en el estudio de los asuntos democráticos y de su evolución. Desde 2008 propuso un marco de evaluación basado en dos principios: el control popular y la igualdad política, al considerarlos como soportes de la vida democrática. Adicionalmente, considera siete valores de mediación (participación, autorización, representación, rendición de cuentas, transparencia, sensibilidad y solidaridad) y cuatro pilares (ciudadanos, ley y derechos, gobierno representativo y responsable, sociedad civil y participación popular, y democracia más allá del Estado). Ha generado índices para 163 países, tomando como referencia distintas fuentes de información, el índice del estado de la democracia es resultado de analizar 116 indicadores desarrollados por académicos y organizaciones a partir de diferentes fuentes de información.

El índice del Estado Global de la Democracia es un índice sintético que evalúa simultáneamente diversos asuntos. Entre sus objetivos se señalan:

- a) Evaluar el estado de la democracia, analizando asuntos relevantes,
- b) diagnosticar las tendencias críticas globales y regionales que reflejan el estado de la democracia en el mundo,
- c) identificar oportunidades para mejorar o reformar la democracia, atendiendo la diversidad, el género y la seguridad,
- d) reconocer las buenas prácticas,
- e) y cumplir el mandato de la organización y de las áreas de *expertise* de IDEA (Skanning, 2020, p. 7).

Como puede apreciarse, el estudio de la democracia, un asunto abstracto e inabarcable, aunque variado y, en ocasiones, extremadamente básico, ha captado un creciente interés mundial con un notorio perfeccionamiento. Diferentes dimensiones de la vida democrática se estudian y sus atributos se observan de manera sistemática. La propuesta de indicadores, escalas, técnicas e índices es considerable, de modo que no se puede estar satisfecho con el mero escrutinio de los votos, ya que lo electoral es apenas una de sus esferas. La evolución temporal y la comparación son focos de atención de sus investigadores cuando reconocen que el detrimento es posible y que el camino democrático es sinuoso y difícil. Adicional a que esta forma de gobierno requiere cuidado y vigilancia permanentes. Aun así, es de advertir que el estado de las iniciativas es germinal y que se requieren técnicas robustas, indicadores más finos, así como mayor investigación. Por ser tan importante para la vida en comunidad, lo político así lo demanda.

9.5 SÍNTESIS CAPITULAR

La evaluación política constituye una modalidad con gran potencial para la vida en sociedad y para el desempeño gubernamental. Sin duda, constituye un desafío intelectual que exige partir de lo abstracto, conceptualizar, operacionalizar, identificar variables y construir indicadores con base en la realidad observada. Ello es, de las ideas del mundo abstracto, hasta llegar a los elementos empíricos (el mundo real y concreto). El evaluador utilizará las técnicas analíticas que le permitan encontrar medidas sintéticas para emitir los respectivos juicios. Con base en estas medidas, establecer comparaciones y estrategias para mejorar los desempeños políticos bajo estudio.

REFERENCIAS

- Bastian, S. y Luckman, R. (2003). *Can democracy be designed? The Politics of Institutional Choice in Conflict-Turn Societies*. Zed books.
- Beetham, D.; Carvalho, E.; Landman, T. y Weir, S. (2009). *Evaluar la calidad de la democracia. Guía práctica*. Instituto Internacional para la Democracia y la Asistencia Electoral, IDEA International.
- Bobbio, N. (1997). *Estado, gobierno y sociedad. Por una teoría general de la política*. Fondo de Cultura Económica.
- Carrillo, J. (2009). La dimensión política de la evaluación de políticas. *Ópera*, 9, 53-65.
- Carrillo, J. y Toca, C. E. (2008). Las nuevas relaciones en la acción pública: roles emergentes en los procesos decisorios. *Papel Político*, 13(1), 95-121.
- Cepal, Comisión Económica para América Latina (2015). Seminario “Latinobarómetro: 20 años de opinión pública latinoamericana”. Consultado el 11 de febrero de 2023. www.cepal.org/fr/node/34337
- Connor, E. C. (2019). *The Social Dimension of Political Values*. Previous versions of this manuscript were presented at MPSA (2017 and 2018) and APSA (2017 and 2018). http://www.elizabethchaseconnors.com/images/Full_Paper_Single.pdf
- Corporación Latinobarómetro (2021). *Informe 2021. Adiós a Macondo*. <https://www.latinobarometro.org/latContents.jsp>
- Douglas, H. E. (2009). *Science, Policy and the Value-free Ideal*. University of Pittsburgh Press.
- Easton, D. (1965A). *A Framework for Political Analysis*. Prentice-Hall Publishers.
- Easton, D. (1965B). *A Systems Analysis of Political Life*. John Wiley and Sons.
- Fisher, F. y Gotweiss, H. (eds.) (2012). *The Argumentative Turn Revisited. Public Policy as a Communicative Practice*. Duke University Press.
- Grace, J. y Grace, P. (2017). Participation Handbook. EU-funded Erasmus+ Programme. <https://360participation.com/wp-content/uploads/2017/11/Participation-Handbook.pdf>
- Howarth, D. y Stavrakakis, Y. (2000). Introducing Discourse Theory and Political Analysis, 1-37, en *Discourse Theory and Political Analysis*, editado por David R. Howarth, Aletta J. Norval y Yannis Stavrakakis. Manchester University Press.
- Howarth, D. (1995). La teoría del discurso, 125-142, en D. Marsh y G. Stoker (eds.), *Teoría y métodos de la ciencia política*. Alianza Editorial.
- Ingram, H. y Schneider, A. L. (2006). Policy Analysis for Democracy, 169-189, en Michael Moran, Martin Rein y Robert Goodin (eds.), *Oxford Handbook of Public Policy*. Oxford University Press.
- Kymlicka, W. y Norman, W. (1994). The Return of the Citizen: A Survey of Recent Work on Citizenship Theory. *Ethics*, 104(2), 352-381.
- Lasswell, H. D. (2000). La orientación hacia las políticas, pp. 79-103, en: Luis Aguilar V. *El estudio de las políticas públicas*. Miguel Ángel Porrúa.
- Majone, G. (1997). *Evidencia, argumentación y persuasión en la hechura de políticas*. México: Fondo de Cultura Económica.

- Meur, S. (2018). *Culture, Values, and Public Policy*. Policy Fellowship Programme, Bath University. <https://www.bath.ac.uk/publications/culture-values-and-public-policy/8>
- Rein, M. (2006). Reframing Problematic Policies, 389-408, en Michael Moran, Martin Rein y Robert Goodin (eds.) *Oxford Handbook of Public Policy*.
- Sartori, G. (1993). *¿Qué es la democracia? Tribunal Federal Electoral*. Patria.
- Sen, A. K. (1999). Democracy as a Universal Value. *Journal of Democracy* 10(3), 3-17. <https://doi.org/10.1353/jod.1999.0055>
- Skaaning, S. E. (2020). *The Global State of Democracy Indices Methodology Conceptualization and Measurement Framework, Version 4*. Instituto Internacional para la Democracia y la Asistencia Electoral, IDEA International. International Institute for Democracy and Electoral Assistance.
- Stewart, J. (ed.). (2009). What Are Policy Values?, 14-32, en: Jenny Stewart (ed.), *Public Policy Values*. Palgrave Macmillan. https://doi.org/10.1057/9780230240759_2
- The Economist Intelligence Unit (2021). The Democracy Index 2020, in *Sickness and in Health?* <https://www.eiu.com/n/campaigns/democracy-index-2020/#:~:text=Democracy%20Index%202020%3A%20In%20sickness,lives%20from%20a%20novel%20coronavirus>
- Van Dijk, T. A. (1977). *Text and Context. Explorations in the Semantic and Pragmatics of Discourse*. Longman.
- Van Dijk, T. A. (1998). *Ideology. A Multidisciplinary Approach*. Sage.

CAPÍTULO 10

Evaluación de impacto

INTRODUCCIÓN

Una concepción adicional sobre cualquier política pública es aquella que advierte que se diseña e implementa con la finalidad última de alterar, cambiar o modificar una situación existente, por lo regular, problemática o adversa. En otros términos, la política pública se diseña para generar un impacto. Por ello, es normal pretender la mejora de los ingresos mensuales, de la condición de salud o del nivel de aprendizaje de los ciudadanos. Frente a estas pretensiones surgen inquietudes como en qué proporción se lograron o qué hubiera ocurrido si no se hubiera implementado la política. Verificar el impacto de una intervención es, seguramente, uno de los aspectos más cruciales y desafiantes en la evaluación de las políticas públicas. En efecto, demanda algunas condiciones especiales para poder presentar evidencia sobre las repercusiones y transformaciones de la condición o el bienestar de los beneficiarios del programa o política. En este sentido, son las agencias gubernamentales y las agencias de cooperación las interesadas en verificar los efectos externos de las distintas intervenciones emprendidas sobre grupos de interés como hogares, comunidades o colectivos sociales. Solo esta verificación puede ofrecer certezas sobre el cumplimiento de sus propias responsabilidades. La inquietud es apenas lógica, ya que se emprenden acciones de política con la idea de afectar las condiciones de vida o los niveles de bienestar, de modo que la sensación sea que los esfuerzos realizados se tradujeron en mejoras y que los esfuerzos valieron la pena.

No se descarta que algún lector haya reparado en la diferencia entre un resultado y un impacto, por tratarse de dos términos que, por lo normal, se confunden e incluso se emplean como sinónimos. Al respecto, es necesario precisar que mientras el resultado equivale al efecto interno de una actividad implementada, el impacto corresponde al efecto producido, pero en el entorno. De hecho, la literatura distingue los indicadores de resultado y los de impacto, por ejemplo, frente a la mejora de la salud del adulto mayor, un indicador de resultado sería el presupuesto 2022 ejecutado para dicho fin, con relación al presupuesto 2021; mientras que el de impacto podría ser la cantidad de adultos mayores atendidos médicamente sobre la totalidad de adultos mayores, demandando atención médica (aún no la

reciben). Retomando el párrafo anterior, el impacto recae sobre los adultos mayores de las familias, las comunidades y los hogares más necesitados. Los presupuestos se asumen como recursos internos, por lo que es de interés centrarse en el impacto sobre los individuos que hacen parte del capital social (adultos mayores).

Este capítulo presenta las principales consideraciones sobre las técnicas utilizadas para determinar el impacto de un programa, intervención o política pública. Para empezar, se presentan los conceptos básicos, las motivaciones y los procedimientos estándares para emprender una evaluación de impacto. A continuación, se reflexiona sobre uno de sus asuntos centrales: los contrafactuales. Para ilustrar sobre los desafíos que encarna el análisis de impacto, se abordan dos modalidades representativas y con enormes potencialidades: la estrategia de las discrepancias en las diferencias y la regresión discontinua. Considerado esta un valioso aporte para aquellos que emprenderán ejercicios de evaluación cuantitativa.

Como práctica evaluativa, la evaluación de impacto exige una importante dosis conceptual. Un énfasis recurrente de los expertos se observa en la necesidad de imprimir rigor a sus procedimientos; en especial, los referidos a los procesos de selección de los grupos de tratamiento y control/comparación. En efecto, la literatura especializada sugiere seguir estrictos procedimientos para elegir estos dos grupos, garantizar comparaciones juiciosas y evaluaciones justas, y aspiraciones explícitas de la mayoría de las técnicas consideradas para tal fin. Esos deseos hacen que no solo los procedimientos de evaluación de impacto sean muy exigentes, sino también que los evaluadores los sigan.

10.1 LA MEDICIÓN DEL IMPACTO

A pesar de los enormes esfuerzos de las agencias supranacionales para asistir a las distintas agencias del desarrollo y para impulsar la evaluación de impacto, en su literatura especializada no logra encontrarse lo que verdaderamente significa evaluar el impacto. Por fortuna, se dispone de algunas referencias que sirven para determinar su significado y ofrecer claridad sobre las técnicas que intentan valorarlo. A este tenor, el manual-guía de la Red de Redes de Evaluación de Impacto (Network of Networks in Impact Evaluation, NONIE) retoma el concepto emanado del Comité de Asistencia al Desarrollo (DAC, por sus siglas en inglés) de la OCDE. Propone así, estudiar el impacto de “los efectos positivos o negativos, primarios o secundarios de largo plazo producidos por una intervención de desarrollo, directa o indirecta, intencional o eventual” (Leeuw y Vaessen, 2009, p. ix).

Complementariamente, los estudios de impacto “son tipos particulares de evaluación que buscan dar respuesta a preguntas específicas de causa-efecto” (Gertler *et al.*, 2016, p. 8) del tipo, ¿cuál es el impacto o efecto causal de un programa sobre un resultado de interés? La evaluación, entonces, está basada en la dimensión causal y su pregunta puede estar formulada en cualquier tipo de contexto o escenario, esto es, donde se desarrolle un proyecto o un progra-

ma sin importar el sector (educativo, salubre, ambiental, etc.). La evaluación de impacto intenta “determinar más ampliamente si el programa tuvo los efectos deseados sobre los individuos, los hogares y las instituciones” (Baker, 2000, p. 1), busca también verificar si dichos efectos son atribuibles al programa de intervención. Asimismo, este tipo de análisis explora las consecuencias no intencionales sobre los beneficiarios, sean estas positivas o negativas.

Una evaluación de impacto analiza los cambios sobre la población objetivo, sean negativos o positivos, intencionales o no, de un proyecto, programa o política, cuantificando las dimensiones de su impacto. La evaluación de impacto establece el efecto causal sobre uno o varios resultados y valorando aquellas transformaciones atribuibles de manera concreta a una decisión (proyecto, programa o política). En este sentido, el efecto es sinónimo de transformación, cambio, alteración, ocurridos o derivados de una intervención vía un proyecto, programa o política. La causalidad opera para determinar si la iniciativa emprendida es la fuente de tales transformaciones o si tuvo un impacto. Por ello, las técnicas que apoyan este tipo de evaluación estudian los efectos causales ocurridos en un periodo (análisis inter-temporal) y buscan dar respuesta a preguntas específicas del tipo causa-efecto.

La evaluación de impacto es un ejercicio de valoración que exige dejar de lado los recursos y los procesos y, en consecuencia, las actividades de monitoreo y control. Se centra decididamente en los resultados ulteriores y en los efectos sobre la vida de los miembros de la comunidad. Esta visión coincide con lo que, por lo regular, se denomina la hechura de las políticas basadas en evidencias (*evidence-based policy making*), cuya intención es mostrar si dichos cambios en las vidas de las personas pueden atribuirse a las intervenciones institucionales (*Swiss Agency for Development and Cooperation, SDC, s/f*). A la par, evaluaciones de este tipo apoyan procesos decisionales que conduzcan a la ampliación o extensión de aquellas iniciativas que demuestren impactos positivos o, en caso contrario, a abandonar aquello que no produzca los impactos deseados. Contribuciones adicionales tienen que ver con el aprendizaje para el mejoramiento y la comparación de las políticas: mejorar el diseño de política y comparar su efectividad. Una adecuada evaluación del impacto permite determinar si los proyectos están funcionando o no.

10.2 PROCESO DE EVALUACIÓN DEL IMPACTO

Los procedimientos estándares de la evaluación de impacto proponen una serie de fases para realizarla de principio a fin. Como lo reiteran los especialistas, lo principal de este tipo de evaluaciones es generar evidencia sobre el desempeño de una política para el uso de las autoridades o de los responsables e interesados en el programa (Gertler *et al.*, 2016). Este tipo de estudios interesa a diversos actores (*stakeholders*): gobernantes, oposición, sociedad civil e, incluso, los propios beneficiarios de la política. La generación de evidencia demanda mucho rigor en los procedimientos investigativos que sustentan la selección de métodos adecuados, de unidades de análisis y de técnicas a aplicar.

Dado que la decisión de evaluar el impacto compromete recursos de distinta naturaleza, conviene establecer si es necesaria, conveniente y útil. En consecuencia, debe verificarse su justificación y condiciones para adelantarla. Una serie de interrogantes pueden orientar la decisión: ¿el programa amerita valoración?, ¿existe evidencia sobre el adecuado funcionamiento del programa?, y ¿hay disponibilidad de recursos necesarios para adelantar la evaluación? La primera pregunta indaga sobre la existencia de un interés supremo, crítico o valioso, sopesando la urgencia, relevancia o extensión del programa. Las demás respuestas orientarán la decisión de una evaluación de impacto que preserve características procedimentales y científicas (Gertler *et al.*, 2016, p. 27).

Otras cualidades consideradas para el proceso de evaluación son:

1. Innovador, esto es, garantizar la utilización de nuevos y promisorios enfoques.
2. Replicable, implica que puedan reproducirse en diferentes contextos y ampliarse a distintas dimensiones.
3. Estratégico y relevante para que su evidencia apoye la toma de decisiones sobre el futuro del programa, ya que está de por medio la asignación de recursos, su expansión o su reforma.
4. Influyente, en relación con el numeral anterior, indicando su utilidad para todos los involucrados.
5. Carecer de evaluación previa.

Toda evaluación de impacto comienza con la respuesta a interrogantes causa-efecto, equivale a evaluar el efecto causal del programa. Los exponentes de las técnicas de evaluación encuentran este procedimiento como el más exigente y recomiendan el mayor rigor para dar respuesta al interrogante ¿para qué se evalúa el impacto?

Aunque simple, este interrogante no es accesorio y es que para conocer la efectividad de una intervención tiene que producirse evidencia, de otra manera cualquier juicio, a favor o en contra, carecerá de validez. Desde el punto de vista investigativo, la evaluación cuantitativa del impacto es un asunto que demanda rigor científico, a tal punto que algunas tesis doctorales han optado por la investigación evaluativa (generación de conocimiento replicable). Sus resultados constituyen un referente objetivo de la toma de decisiones por parte de los responsables e involucrados de una intervención. Pudiera afirmarse, entonces, que la evaluación de impacto promueve la responsabilidad, las buenas prácticas y, por supuesto, el uso eficiente de los recursos públicos o privados. Las políticas públicas como campo disciplinar pueden beneficiarse sustantivamente, en la medida que determinen aquello que funciona o no de un programa particular. A continuación, se enlistan asuntos clave para la concepción de un ejercicio de evaluación de impacto.

- Identificar las preguntas de evaluación base del proceso evaluativo y orientadores de la acción.
- Disponer de teoría del programa, fundamento teórico base para la intervención.
- Comparar grupalmente, generando bases de contrafactuales creíbles.
- Triangular los hallazgos.
- Contextualizar el ejercicio.

Adicional, toda evaluación del impacto debe considerar modalidades prospectivas y retrospectivas (Gertler *et al.*, 2016). Las prospectivas se aplican al mismo tiempo que el diseño del programa y los datos de la línea base para los grupos de tratamiento y de control o intervención son recolectados antes de su implementación. Por el contrario, las retrospectivas se aplican después de la implementación del programa, determinando *ex post* los grupos de tratamiento e intervención. Las evaluaciones prospectivas tienen una mayor probabilidad de generar resultados fuertes y creíbles, por tres razones básicas: la información de la línea base puede establecerse con más precisión e interés, previo al inicio del programa o política; al definir, en la fase de planeación, medidas de éxito del programa, hay una mayor conciencia de lo que se quiere alcanzar; y elección de los grupos de tratamiento y comparación antes de la implementación del programa. En los países del Tercer Mundo no hay tiempo para la planeación de la evaluación del impacto de las acciones gubernamentales; en ocasiones, lo único que puede anticiparse es el fracaso del ejercicio evaluativo.

10.3 EL ASUNTO DE LOS CONTRAFCTUALES

Uno de los asuntos cruciales en la evaluación de impacto son los contrafactuales sobre los que descansan unos procedimientos y la aplicación de técnicas. Estimar la efectividad de una política en curso mediante la comparación de los resultados individuales de los responsables de la política, es un propósito muy simple de entender, pero difícil de llevar a la práctica. A esta idea hipotética de la comparación de resultados o impactos se le denomina contrafactual. Esta simultaneidad de la medición de los resultados o impactos sobre un individuo o grupo, en presencia y en ausencia de una intervención, constituye la condición principal de la evaluación de impacto. El contrafactual raramente se observa en la realidad, al resultar imposible que, en paralelo, un individuo o grupo se beneficien y no se beneficien de una política. La medición simultánea de la incidencia no es posible, pues a todas luces constituye un reto, representado precisamente en el contrafactual.

El desafío de la evaluación de impacto es encontrar un individuo o grupo que sea tan similar como sea posible y que no participen en el programa para conformar el grupo de control o referente *proxi* para la comparación. En buena parte, la calidad de la evaluación depen-

de de hallar este grupo de comparación para estimar la situación de los participantes en ausencia del programa (Gertler *et al.*, 2016). Sin duda, disponer del grupo contrafactual es central para la evaluación de impacto (Baker, 2000). Por tanto, para cumplir este requisito se dispone de dos tipos de diseños: experimental (aleatorio) y cuasi-experimental (no aleatorio). Aun así, se reconoce que el procedimiento puede estar afectado por la historia, el sesgo de selección y la contaminación (mayores detalles pueden consultarse en Baker, 2000).

“Qué hubiera pasado si...”, es la reflexión que permite iniciar una evaluación de impacto y el proceso de análisis comparativo de dos grupos distintos para establecer el diferencial derivado de la intervención (política pública): el grupo de tratamiento y el grupo de control. Se puede concebir, entonces, el impacto de un programa como la diferencia en resultados (Y) para una misma unidad de análisis (persona, hogar, comunidad, etc.) en presencia o ausencia de un programa o política. Dada la inferencia causal y los contrafactuales, medir la misma unidad en dos diferentes estados al mismo tiempo es imposible, por ello se reconoce como el problema del contrafactual: cómo medir lo que hubiera pasado si otras circunstancias hubieran prevalecido. Esta idea podría asimilarse a los *ex futuros*; propuesta esbozada, pero no desarrollada por Miguel de Unamuno e invocada por Héctor Abad Faciolince (2009) y quien afirma que la literatura es la que analiza las variantes de la vida. La literatura es “una manera de lidiar con nuestros *ex futuros*: con eso que no somos, pero que podríamos llegar a ser o que pudimos haber sido” (Abad, 2009, p. 254). Los contrafactuales, entonces, superan esa deficiencia de no tenerse clones o copias genéticas y que podría ser asimilado al ejercicio mental y tormentoso, que consiste en pensarse, no como se es, sino como podría haberse sido (Abad, 2009). Entonces, el pensamiento contrafáctico o contrafactual descansa en la suposición de mundos alternativos a considerar en determinados momentos. Reflexionar acerca de lo que hubiera podido haber ocurrido si los acontecimientos hubieran sido otros, al parecer es un ejercicio común del ser humano. Según Pelegrín (2010), es una manera de construir alternativas a nuestro pasado y, en última instancia, a nuestro presente; el interés por el pensamiento contrafáctico surgió en la filosofía en la década de los años cuarenta y se extendió a otras disciplinas, especialmente a la psicología. Las historias alternativas sirven de base para la evaluación de impacto cuando se plantea el problema del contrafactual o de cómo medir lo que hubiera pasado si otras hubieran sido las circunstancias.

Como ya se advirtió, el primer desafío consiste en establecer dos grupos: uno de tratamiento y otro de control, cuya comparación permita establecer el efecto del programa. Asimismo, se trata de garantizar grupos idénticos estadísticamente para estimar el resultado contrafactual o el que hubiera prevalecido para el grupo de tratamiento en ausencia de este. El uso de variables dicotómicas o también conocidas como variables *dummy* [0, 1] resulta conveniente para estudiar este diferencial en los resultados. La estimación del impacto se soporta en la siguiente fórmula:

$$I = (Y|P = 1) - (Y|P = 0) \quad (10.1)$$

La estimación del impacto, I , es el diferencial entre dos resultados, cuando se participa en un programa y cuando no. Esta fórmula establece que el impacto de un programa, P , sobre un resultado (Y) es la diferencia entre el resultado con el programa (cuando $P = 1$) y el mismo resultado (Y) sin el programa (cuando $P = 0$). $P = 0$ representaría lo que hubiera sucedido en el resultado, si la persona/unidad analizada no hubiera participado del programa (ausencia).

En una evaluación de impacto es relativamente sencillo obtener el primer término de la fórmula básica ($Y|P = 1$), el resultado con un programa (bajo tratamiento), basta con medir el resultado de interés para el participante del programa. Sin embargo, el segundo término de la fórmula ($Y|P = 0$) no se puede observar para el participante, es así que se debe estimar el contrafactual. La validez interna de la evaluación estará garantizada por este grado de comparabilidad entre los dos grupos y será el pilar para establecer el impacto causal del programa (Gertler *et al.*, 2016). El contrafactual, sin duda, es clave para la evaluación del impacto y, por suerte, las propuestas sugeridas para su determinación intentan en esencia preservar el rigor y la validez estadística, de modo que los dos grupos configurados sean comparables y la estimación del impacto sea más precisa. Así, por ejemplo, la selección aleatoria de los participantes (*Random Controlled Trial*, RCT), en esencia, tiene que ver con la decisión aleatoria de aquellos beneficiarios del programa que harán parte del estudio. Los promotores de esta modalidad no solo encuentran una selección justa de los beneficiarios, sino también la manera adecuada para seleccionar el contrafactual y desarrollar la evaluación de impacto (Gertler *et al.*, 2016). Es bien sabido que, en la práctica, raramente se considera la selección de beneficiarios mediante sorteos (loterías) u otros mecanismos, en su lugar dominan los intereses partidistas, la cooptación y todas aquellas prácticas que caracterizan la realidad política.

10.4 TÉCNICAS PARA LA EVALUACIÓN DE IMPACTO

En la actualidad se disponen de varias posibilidades para el análisis del impacto como los métodos experimentales y los cuasi-experimentales. Cada posibilidad tiene sus propios desarrollos, por lo que es preciso conocer sus restricciones para decidir su aplicabilidad en un ejercicio evaluativo. Los métodos aleatorios exigen que la selección de los individuos pertenecientes a los grupos de tratamiento y control (comparación) se efectúe al azar. La única diferencia que prevalece entre los individuos es el acceso a los beneficios del programa evaluado. Desde luego, también se reconocen las diferencias derivadas de los errores de muestreo, del tamaño de las muestras de los grupos, etc. Los métodos cuasi-experimentales son elegidos cuando la conformación de los grupos de tratamiento y control no pueden conformarse de manera aleatoria. Por lo regular, los grupos se conforman después de la intervención, aunque también pueden hacerse antes, pero la selección de los participantes no es aleatoria. Estos métodos cuasi-experimentales proporcionan alguna ventaja al diseñador de la evaluación por resultar más rápidos y económicos de implementar (Baker, 2000). Aun así,

se reconocen algunas desventajas como: la pérdida de robustez estadística de la técnica, el aumento en la complejidad de los procesos, y el sesgo de la selección que es un problema emergente. Dentro de los cuasi-experimentales destacan: el método de emparejamiento o controles construidos, el método de las variables instrumentales y el método de las comparaciones reflexivas. El método de las dobles diferencias (la diferencia en las diferencias) y el diseño de la regresión discontinua son dos posibilidades analíticas elegidas para ilustrar la manera de proceder en la evaluación de impacto.

10.4.1 El método de la diferencia en las diferencias

Cuando no es posible garantizar una elección, selección justa o aleatoria, se recomienda utilizar otros métodos diferentes de evaluación. El método de las diferencias en las diferencias (DeD) compara los cambios en los resultados a través del tiempo entre el grupo de tratamiento (beneficiarios del programa) y el grupo de comparación (no beneficiarios del programa). Se trata de establecer dos diferencias:

- a) La calculada sobre los beneficiarios bajo las condiciones imperantes antes y después; los factores imperantes son constantes en el tiempo dado que se están comparando las mismas personas a lo largo del tiempo.
- b) Aquella que captura la variación de los factores al medir las condiciones antes y después de los no beneficiarios, pero expuestos al mismo conjunto de condiciones ambientales. La tabla 10.1 presenta la situación de la evaluación de impacto para los dos grupos bajo estudio.

TABLA 10.1. EVALUACIÓN DEL IMPACTO COMPARANDO DOS GRUPOS

Grupo	Condición inicial	Condición final
Beneficiarios (participantes)	A	B
No beneficiarios (no participantes)	C	D
Primera diferencia	B-A	
Segunda diferencia	D-C	

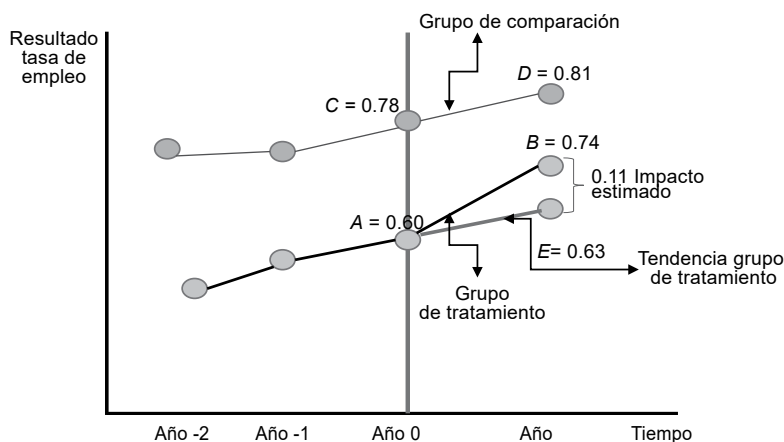
Fuente: elaboración propia con base en Gertler *et al.*, 2016.

Para ilustrar el procedimiento, Gertler *et al.* (2016), abordan un programa de reparación de caminos. Para mejorar las condiciones de empleo de un distrito (área geográfica que puede ser rural o urbana) se propone un programa de mantenimiento de caminos, que proporcionaría oportunidades de trabajo a los habitantes de la comunidad y determina el grupo de tratamiento. El método de las diferencias en la diferencia requiere establecer las tasas de empleo de

la comunidad previas y posteriores a la implementación del programa, determinando así la primera diferencia. A la par, se selecciona otro distrito a donde no llegó el programa, de preferencia con gran similitud con las condiciones imperantes en el grupo de tratamiento, midiéndose las tasas de empleo cuando el programa empieza y cuando este termine. Este es el grupo de control y los niveles de empleo, antes y después, constituyen la segunda diferencia.

La idea central del método de las diferencias en las diferencias es comparar los cambios antes y después de la implementación, tomando como referencia los individuos que viven en el distrito que fueron enrolados en el programa (grupo de tratamiento) contra los cambios en el empleo en distritos similares no beneficiarios (grupo de comparación).

En un plano cartesiano (gráfica 10.1), el eje horizontal representa la línea de tiempo, marcando el antes y el después del programa (línea roja perpendicular). Por otra parte, el eje vertical representa la tasa de empleo del grupo de tratamiento y del grupo de comparación, en los momentos respectivos (año -2 , -1 , 0 , 1 , $2 \dots$). Al inicio del programa (año 0), las tasas de ocupación en el grupo de tratamiento eran de 0.60 y de 0.78 en el grupo de comparación. Estas se representan con A y C , respectivamente. De no emprender el programa, la tasa para el grupo de tratamiento hubiera sido de 0.63 , representada por la línea de tendencia.¹ Al implementar el programa se alcanzó un nivel de 0.74 (B). La tendencia del grupo de comparación llegó a un nivel de 0.81 (D), sin importar las razones por las cuales se obtuvo. El impacto estimado, entonces, es la diferencia entre la tasa de empleo en el año 1 (0.74) y la tendencia en ausencia del programa (0.63).



GRÁFICA 10.1. DIFERENCIAS EN LAS DIFERENCIAS DEL PROGRAMA REPARACIÓN DE CAMINOS

Fuente: elaborado a partir de Gertler *et al.*, 2006, p. 132.

¹ No sobra recordar que la ecuación de la línea de tendencia puede calcularse mediante el análisis de regresión lineal simple y el nivel del año 1 se establece mediante un pronóstico, utilizando la ecuación de regresión para el año 1.

El método propone el cálculo de las siguientes diferencias:

1. De los resultados del grupo de tratamiento (Y), antes y después ($B - A$).
2. De los resultados del grupo de comparación (Y), antes y después ($D - C$).
3. Entre las diferencias (DD) de los resultados para el grupo de tratamiento ($B - A$) y para el grupo de comparación ($D - C$). Este es el impacto esperado $(B - A) - (D - C)$.

Alternativamente, se propone el cálculo de las siguientes diferencias:

1. Entre los grupos de tratamiento y comparación después de la intervención.
2. Entre los grupos de tratamiento y comparación en el momento anterior a la intervención.
3. Entre 1 y 2.

$$\begin{aligned} \text{Impacto estimado} = & (B - D) - (A - C) \\ & (0.74 - 0.81) - (0.60 - 0.78) \\ & (-0.7) - (-0.18) \\ & 0.11 \end{aligned}$$

Puesto de esta manera, el impacto estimado (0.11) sería la diferencia vertical del después de los dos grupos (0.74 - 0.81) y, la diferencia vertical del antes de los grupos de tratamiento y comparación (0.60 - 0.78). Reemplazando con los datos del ejercicio se obtiene la tabla 10.2.

TABLA 10.2. LA DIFERENCIA ENTRE LAS DIFERENCIAS

Grupo	Después	Antes	Diferencia
De tratamiento	B	A	$B - A$
	0.74	0.60	0.14
De comparación	D	C	$D - C$
	0.81	0.78	0.03
Diferencia	$B - D$	$A - C$	$(B - A) - (D - C)$
	-0.7	-0.18	0.11

Fuente: elaboración propia.

La gráfica 10.1 y la tabla 10.2, además de comprobar las dos posibilidades para establecer el impacto estimado, revelan que el impacto ocurre sobre la variable dependiente (y), tasa de empleo estudiada en dos momentos antes y después de los beneficios de una política. Una lectura horizontal conlleva a la diferencia temporal del antes y el después de los dos grupos bajo estudio, y a la diferencia de los dos grupos bajo estudio en cada uno de los momentos.

10.4.2 Método de regresión discontinua

Durante las últimas décadas el Diseño de la Regresión Discontinua (Regression Discontinuity Design, RDD) se ha convertido en el método idóneo para la evaluación del impacto de un programa o política. A pesar de popularizarse su uso, su comprensión para la aplicación aún representa un desafío especialmente en contextos latinoamericanos. Reconocido como método cuasi-experimental que facilita o supera, de alguna manera, las dificultades que encarna la elección de los contrafactuales, las posibilidades de aplicación deben reconocerse. Tres componentes han sido propuestos para tal fin: un puntaje obtenido, un tratamiento y un punto de corte (Cattaneo, Idrobo y Titiunik, 2019). Si no se cumplen estas condiciones, el método no procede. Los autores mencionados consideran que la regresión discontinua es un diseño de investigación más que una estrategia de estimación y, en consecuencia, los puntajes, el tratamiento y el punto de corte deben existir y estar bien definidos y las relaciones entre ellos deben satisfacer condiciones particulares: objetivas y verificables. Para la regresión discontinua, RD, debe garantizarse:

1. Un puntaje o variable continua (X_i), denominada en la literatura como la variable en ejecución, la de asignación o forzada (*running variable, assignment variable o forcing variable*).
2. Un tratamiento, previamente determinado por una regla conocida y que se identifica como variable dicotómica: (1) si obtuvo el tratamiento y (0) si no fue sujeto del tratamiento.
3. Un punto de corte que establece la frontera entre quienes reciben o no el tratamiento, hace referencia a un umbral que representa una frontera que separa los beneficiarios y los no beneficiarios. Se representa como una línea paralela al eje de las X y se considera que es una condición exógena.

Lo anterior indica que la mayoría de los programas sociales tienen algunos requisitos acreditados por los beneficiarios o interesados sujetos de la acción. Los umbrales o límites pueden ser de distinta naturaleza como la edad, el ingreso, el nivel escolar, etc., mismos que pueden ser únicos o concurrir con otros. Estos requisitos se conciben como mecanismos para focalizar

o excluir y atender solo a aquellos que verdaderamente lo requieran. La virtud de la regresión discontinua es que determina si un tratamiento o política es efectiva.

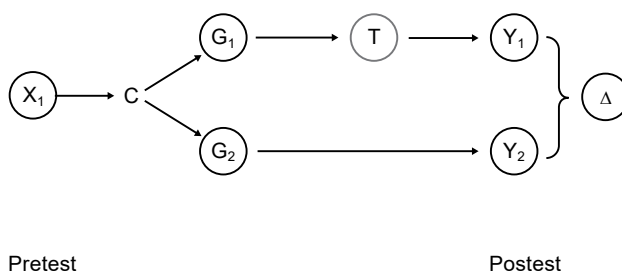
La RD estima el impacto en torno al punto de elegibilidad (umbral o límite), estableciendo la diferencia promedio de los resultados de quienes son beneficiarios (grupo de tratamiento) y quienes no son sujetos del programa o política (grupo de comparación). A este tenor se asume que quienes están muy cerca del umbral, a la izquierda o la derecha, tienen similares características y la única diferencia es el efecto del programa. Esta es una manera no experimental de elegir los contrafactuales, al suponer que aquellos que están en la proximidad del umbral, guardan gran similitud. Bajo estas consideraciones, un estudiante que accede a una beca porque con 8.1 acreditó el requisito del promedio (> 8.0), es muy similar a aquel que se quedó fuera porque su calificación fue 7.9. Mientras que el primero es elegible por superar el límite establecido por las reglas de operación del programa, el segundo no lo superó. El 8.0 es un criterio de elegibilidad.

Por tanto, no sobra precisar que el diseño de RD se decide cuando los beneficiarios son seleccionados porque cumplen estrictamente con los requisitos. Una característica importante de los criterios de elegibilidad es que, por lo regular, se asocian a, o representan en una escala continua, garantizando la posibilidad de hacer un ordenamiento de los puntajes de los candidatos al programa. Los índices de pobreza, por ejemplo, permiten ordenar las familias del menor al mayor en una determinada área geográfica. El diseñador de la política establece el punto de corte y este no debe manipularse políticamente. Así, por ejemplo, el requisito de la edad para obtener una pensión (65 años o más), determina claramente a los beneficiarios; el puntaje mínimo de un examen de ingreso discrimina a los elegibles para estudiar un programa; el nivel salarial es un referente para elegir a quienes demandan capacitación. A continuación, algunos ejemplos de los programas sociales latinoamericanos.

Colombia: “Ser Pilo Paga” otorga recursos a estudiantes que obtengan un puntaje igual o superior a 318 en las pruebas SABER (pruebas de Estado), posibilitándoles su acceso a universidades acreditadas de alta calidad (incluidas las privadas). Este puntaje establece el límite o umbral para quienes pueden ser beneficiarios del programa. Chile: “Me conecto para aprender” entrega computadores a estudiantes matriculados (jóvenes o adultos) en séptimo básico o tercer nivel de educación básica de establecimientos públicos, pero durante determinado periodo del año (fecha). Argentina: “Materno infantil” contribuye al crecimiento y desarrollo adecuado e integral de los niños entre 0 y 5 años. Brasil: “Bolsa Familia” beneficia las familias en situación de pobreza cuyos ingresos mensuales sean menores a 154 reales por persona (unos 40 dólares a octubre de 2015). Al respecto, el Banco Interamericano de Desarrollo precisa que el programa considera una familia como aquella unidad nuclear que, en ocasiones, incluye a otros individuos con o sin parentesco que comparten la vivienda. Los individuos que no cumplan con el requisito mínimo (318 o más puntos, matriculado en séptimo básico o tercer nivel en escuela pública en determinada fecha, menores de cinco años o con ingresos superiores a 154 reales) no serán beneficiarios de los programas en cuestión.

Supuestos de la regresión discontinua

El análisis de regresión discontinua estima el impacto de la selección de beneficiarios de una política pública por el hecho de estar por encima o debajo o a la izquierda o a la derecha de un umbral o punto de corte (Jacob y Zhu, 2012). Como se advirtió, este análisis, en algunas latitudes, ha mostrado cierta popularidad; sin embargo, por sus particularidades, alto contenido técnico y lenguaje especializado, solo una audiencia especializada lo aplica, permaneciendo como una opción exclusiva de ciertos círculos académicos. Donald L. Thistlethwaite y Donald T. Campbell (1960), son reconocidos como los responsables de introducir esta perspectiva para la evaluación de programas sociales. La gráfica 10.2 representa el proceso a surtir para el análisis de regresión discontinua, incluyendo los elementos constitutivos en una secuencia temporal.



GRÁFICA 10.2. PROCESO DE REGRESIÓN DISCONTINUA

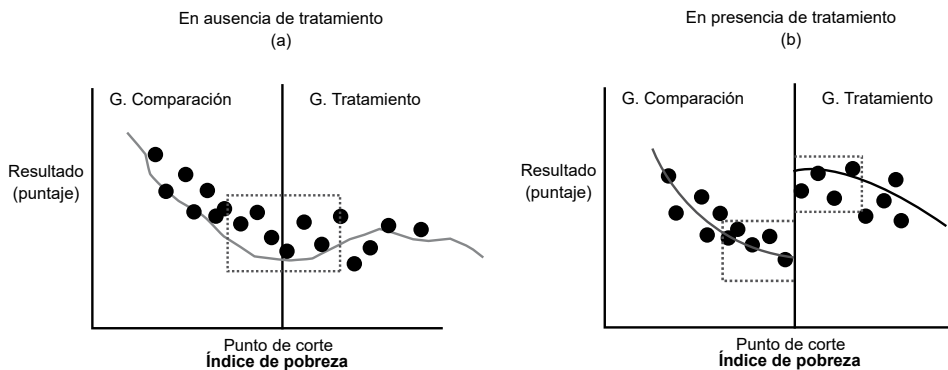
Fuente: elaboración propia.

El análisis inicia con la variable X_1 que representa una característica que ostentan los participantes del programa: puntaje, nivel salarial, edad, etc., representarían dicha variable. Este puntaje, al considerarse una prueba de entrada, es reconocido como pretest y X_1 se denomina variable de asignación (*running variable*). Los resultados obtenidos por los beneficiarios del programa deben compararse con un criterio de elegibilidad, un punto de corte (C) establecido con antelación, el cual es una condición exógena cuyo cumplimiento o no, condiciona la modalidad a utilizarse en la regresión discontinua (*sharp* o *fuzzy*). La aplicación de tal criterio genera dos grupos, uno arriba del criterio y otro abajo. El grupo de tratamiento (G_1) recibirá un tratamiento (T) que asume distintas modalidades, dependiendo de la finalidad u objetivo propuesto: una capacitación, una tutoría, una beca, etc. El grupo de comparación (G_2) no recibirá el tratamiento y estará excluido por el criterio establecido. Una vez implementado T , se aplicará una prueba o se verificará una condición alcanzada, esta etapa será posttest y los puntajes obtenidos generarán las variables Y_1 y Y_2 , o la variable de resultado para ambos grupos (tratamiento y comparación). La regresión discontinua busca conocer el impacto del tratamiento mediante el análisis de Y . Delta (Δ), por su parte, representa la diferencia en los resultados de los dos grupos: el efecto o la efectividad del tratamiento implementado.

La aplicación de tal tratamiento en el punto de corte genera una discontinuidad que permite asumir la existencia de un efecto en el comportamiento de las observaciones que solo puede atribuirse al tratamiento al que fue sometido dicho grupo. El propósito de la regresión discontinua, en consecuencia, es medir la magnitud del efecto. Para otorgar mayor relevancia a la técnica y seleccionar adecuadamente los contrafactuales, se establece una banda de ancho (*bandwidth*) con diferentes medidas en torno al punto de corte, pues se asume que allí los casos solo pueden diferenciarse por el tratamiento recibido. Este procedimiento se denomina aleatorización local (*local randomization*).

Una representación de la relación que existiría entre dos variables en dos momentos (antes y después de un tratamiento) ayuda a entender el proceso (gráfica 10.3). La sección (a) de la gráfica representa en y , el resultado promedio de un examen y cuyos resultados los clasifica como potenciales candidatos a un tratamiento (beca, tutoría, etc.); y en x su nivel de pobreza con un punto de corte o umbral (línea vertical), antes o en ausencia del tratamiento. Se aprecia una relación inversa entre el resultado del examen y el nivel de pobreza, disminuyendo de izquierda a derecha, indicando que el promedio de los resultados disminuye a medida que el nivel de pobreza crece. A mayor pobreza, el promedio de los resultados baja. Tal relación cruza la línea del umbral, indicando que no hay diferencia en los resultados de quienes se ubican arriba o abajo de dicho punto de corte.

La sección (b) de la gráfica representa la situación posterior al tratamiento, mostrando lo que ocurriría en los resultados de una prueba. La beca produce mejores resultados en los estudiantes participantes y por ello se observa un nivel más alto en sus puntajes. En este caso, se observa una discontinuidad, es decir, un salto en el punto de corte en la relación de las dos variables. La dirección y magnitud de dicho salto es una medida directa del efecto causal del tratamiento, en el punto de corte, sobre los resultados de quienes se sometieron al tratamiento.



GRÁFICA 10.3. EFECTO DEL TRATAMIENTO ANTES Y DESPUÉS DEL TRATAMIENTO

Fuente: elaboración propia.

En adición, dos modalidades son consideradas para el análisis de regresión discontinua: el diseño preciso, escrupuloso o nítido (*sharp*), bajo el que todos los elegidos para el tratamiento cumplen estrictamente una condición establecida; y el diseño difuso (*fuzzy*), en el cual algunos participantes no cumplen la mencionada condición.

10.5 APLICACIÓN PRÁCTICA

La elección de la modalidad de la regresión discontinua es responsabilidad del analista en la medida que se cumplan los requisitos establecidos; para un diseño de esta naturaleza se sugiere el siguiente procedimiento (Heiss, 2020):

1. Determinar si la elección de los beneficiarios está basada en algún criterio o regla de asignación.
2. Decidir la modalidad: *fuzzy* o *sharp*.
3. Verificar la discontinuidad de la variable de asignación en el punto de corte.
4. Comprobar la discontinuidad de la variable resultado (dependiente) a través de la variable de selección (*running variable*).
5. Medir la dimensión del impacto o efecto.
6. Establecer las comparaciones del caso.

Para confirmar este procedimiento, Heiss (2020) presenta el siguiente ejemplo hipotético: 1,000 estudiantes presentan un examen al inicio del año escolar, el puntaje obtenido (entre 0 y 100 puntos) por cada estudiante constituye un referente para la asignación de una tutoría (tratamiento). Aquellos que tienen un puntaje igual o inferior a 70, automáticamente son enrolados en un programa de tutoría gratuito impartido en la escuela, orientado a mejorar el desempeño escolar. Al terminar el año presentarán un examen final para verificar su nivel de aprendizaje. La tabla 10.3 presenta las características del archivo de datos.

TABLA 10.3. ENTRADA DE DATOS

Código estudiante	Puntaje (0-100)		Tutoría
	Inicio año	Final año	
			0 = No recibida
			1 = Recibida

Fuente: elaboración propia con datos de Heiss, 2020.

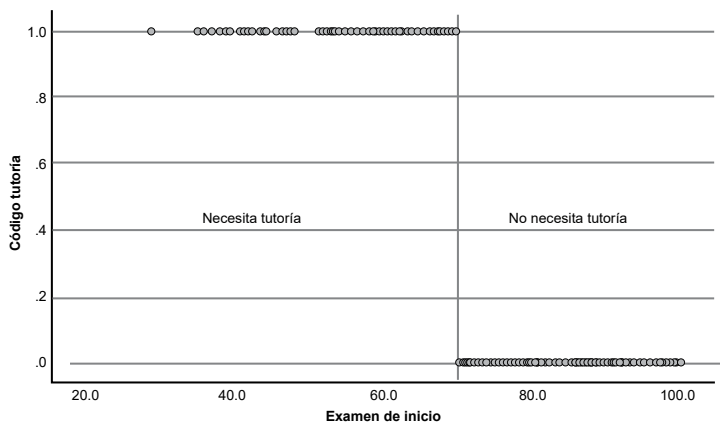
Los seis pasos arriba mencionados se siguen para determinar la efectividad (impacto) del programa de tutoría.

Paso 1. Determinar si la elección de los beneficiarios está basada en algún criterio o regla de asignación.

En efecto, el programa establece un umbral para quienes podrían recibir la tutoría; asimismo, el programa está basado en la siguiente regla de selección:

Punto de corte: 70 puntos = 0 < \Rightarrow Pasan a tutoría
 70 puntos > \Rightarrow No elegibles para tutoría

La gráfica 10.4 muestra la probabilidad de ser elegido para la tutoría programada para estudiantes que no alcanzan 70 puntos en el examen de inicio.² Un diseño *sharp*, solamente elegiría a aquellos estudiantes que están por debajo del punto (línea de corte) y su probabilidad de elección es igual a 1. El eje x representa la variable continua de una escala de calificaciones posibles (0-100), mientras que el eje y muestra la probabilidad (0-1) de ser elegido para el programa. Para el ejemplo de Heiss (2020), 759 estudiantes lograron más de 70 puntos, mientras que 241 resultaron elegibles para recibir la tutoría.



GRÁFICA 10.4. PROBABILIDAD DE SELECCIÓN DE CASOS

Fuente: elaboración propia con datos de Heiss, 2020.

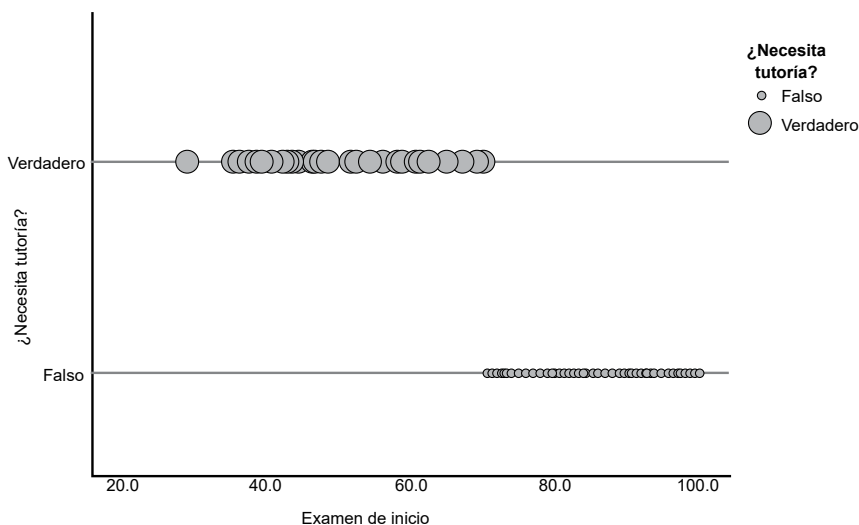
² Esta es una gráfica de dispersión la cual considera una variable continua (puntaje obtenido) en el eje horizontal y una variable dicotómica (0, 1) en el eje vertical. Su utilidad radica en que se puede apreciar la probabilidad de ser elegido para ser parte de un tratamiento (tomar la tutoría). También es ilustrativa porque señala el umbral mediante la línea vertical establecida como requisito en los 70 puntos.

Esta probabilidad $(0, 1)$ de elección al programa determina dos grupos claramente definidos: quienes participan y quienes no participan (beneficiarios y no beneficiarios). Se establecen así, los dos grupos: el de tratamiento y el de control. La elección de estos grupos es condición importante para el diseño de un análisis discontinuo, por lo que no procede la manipulación para su conformación, ello si se quiere garantizar un resultado riguroso. Cuando existe la posibilidad de las excepciones debe considerarse un diseño *fuzzy*, incluyendo algunos casos que superaron el límite o punto de corte.

Paso 2. Decidir la modalidad: *fuzzy* o *sharp*.

Para el caso bajo consideración, se elige el diseño *sharp*, ya que no se consideraron casos con puntuación superior del punto de corte (70 puntos) para la tutoría, es decir, se observaron los lineamientos establecidos. A la par, es importante verificar si alguien hizo trampa o si se cometió algún error estadístico; por ejemplo, elegir a algún estudiante con puntaje superior a 70 como receptor de la tutoría o dejar por fuera de la tutoría a alguien con un puntaje inferior o igual al punto de corte.

La gráfica 10.5 permite confirmar esto, de modo que el puntaje del examen de entrada se representa en el eje horizontal, mientras que el código de tutoría (variable de tratamiento, *false/true*) ocupa el eje vertical. Se aprecia que el umbral establecido en 70 puntos no es traspasado por ningún caso (alumno), lo cual confirma que el diseño *sharp* es el apropiado para el análisis.

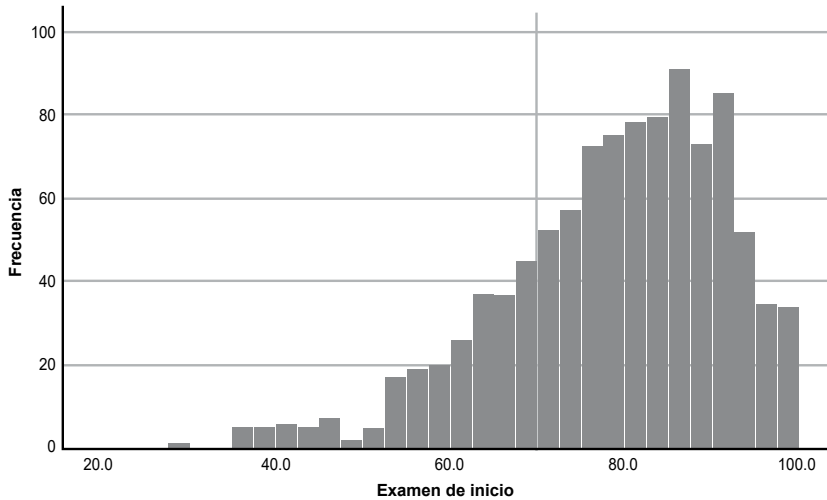


GRÁFICA 10.5. VERIFICACIÓN DE AUSENCIA DE ERRORES DE SELECCIÓN

Fuente: elaboración propia con datos de Heiss, 2020.

Paso 3. Verificar la discontinuidad de la variable de asignación en el punto de corte.

En la variable de asignación de un histograma es posible verificar la presencia de grandes saltos alrededor del punto de corte. La gráfica 10.6, en efecto, no revela mayores diferencias que pudieran indicar alguna manipulación de quienes quisieran estar arriba/abajo del umbral de manera intencional.



GRÁFICA 10.6. DISCONTINUIDAD DE LA VARIABLE DE ASIGNACIÓN

Fuente: elaboración propia con datos de Heiss, 2020.

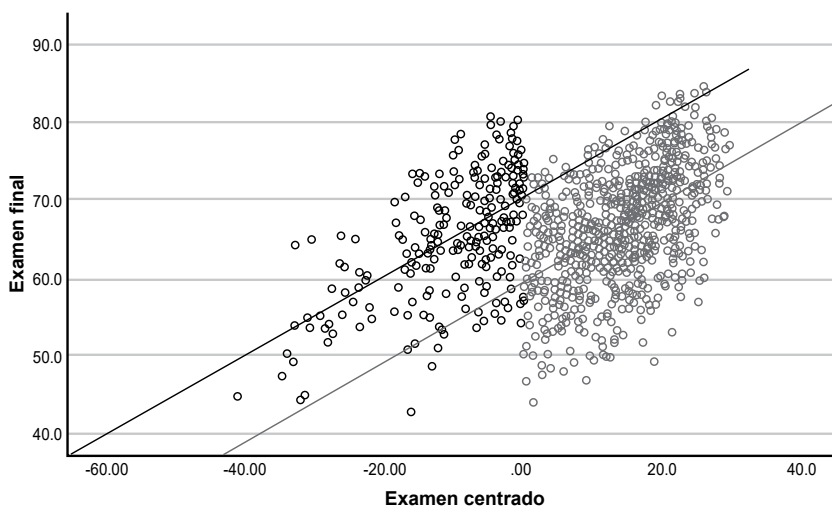
Complementariamente puede trazarse una gráfica de la densidad de la variable de asignación para determinar la continuidad entre quienes están en el programa de tutoría y quienes no. Con este tipo de representaciones el analista puede verificar la idoneidad del diseño de evaluación.

Paso 4. Comprobar la discontinuidad de la variable resultado (dependiente) a través de la variable de selección (*running variable*).

Una gráfica de dispersión entre la variable de asignación (eje horizontal) y la variable de resultado del examen final (eje vertical), podría revelar la discontinuidad a través de las líneas de tendencia de cada grupo (gráfica 10.7). Para efectuar esta prueba es necesario calcular una nueva variable de asignación que se denominará examen de inicio “centrado” que resulta al tomar cada puntaje obtenido por cada estudiante, restándole el puntaje límite o umbral ($x_i - 70$). De esta manera, se observan puntajes superiores e inferiores de este umbral. Son, en realidad, puntajes netos o distancias horizontales al punto de corte valores negativos para quienes son sujetos de tomar la tutoría y valores positivos para quienes no la necesitan. El

umbral a partir de este momento asume el valor cero y se constituye en el límite de los dos grupos de estudio. La nueva variable, “examen centrado”, se representará en el eje horizontal de la gráfica de dispersión mientras que el eje vertical representa el examen final, esto es, después que los alumnos toman la tutoría. A fin de comprender bien el proceso, no sobra advertir la dimensión temporal que existe en este tipo de evaluaciones, retomado las denominaciones pretest y posttest de la gráfica 10.2.

La gráfica 10.7 de dispersión del examen centrado y el examen final constituyen el foco del análisis de la regresión discontinua. Para el ejemplo en cuestión es necesario advertir que las líneas de tendencia de los dos grupos (tratamiento y control) presentan distintos puntos de corte en la línea perpendicular que representa el umbral de los 70 puntos. Los resultados obtenidos por los estudiantes del grupo de tratamiento (izquierda), quienes tomaron la tutoría, presentan un punto de corte más alto (línea de tendencia superior), reflejando el mayor puntaje obtenido de aquellos que no tomaron la tutoría y quienes representan los estudiantes del grupo de control (línea de tendencia inferior). Esta diferencia vertical indica un efecto o un impacto que se aduce únicamente al programa de tutoría que fue impartido.



GRÁFICA 10.7. TENDENCIAS DE LOS GRUPOS DE CONTROL Y TRATAMIENTO

Fuente: elaboración propia con datos de Heiss, 2020.

Paso 5. Medir la dimensión del impacto o efecto.

Como las tendencias lineales establecen puntos de corte para cada grupo, es necesario establecer la magnitud de la diferencia vertical, la cual representa el impacto del tratamiento. Para ello, se utiliza un análisis de regresión múltiple considerando dos diferentes situaciones, a saber: un análisis de regresión considerando todas las observaciones bajo estudio y un aná-

lisis de regresión con las observaciones determinadas por distintos anchos de banda, establecidos, de modo discrecional, alrededor del punto de corte. Para el presente ejercicio se presentan dos bandas de 10 y 5 puntos, respectivamente. De esta manera se aplica un análisis de regresión entre la zona determinada entre 60-80 puntos y entre 65 y 75.

El modelo general de regresión múltiple a estudiar es el siguiente:

$$y_i = \beta_0 + \beta_1 (x_i - c) + \beta_2 x_i + \varepsilon \quad (10.2)$$

Donde:

- y_i = Variable de los resultados del examen final
- $x_0 - c$ = Variable de asignación centrada
- x_i = Variable dictómica (0,1) que indica si participó en la tutoría (1)
- ε = Residuales

El modelo puede representarse, de igual forma, mediante la siguiente ecuación:

$$\text{Examen final} = \beta_0 + \beta_1 \text{examen de entrada centrado} + \beta_2 \text{tutoría} + \varepsilon$$

Análisis de regresión múltiple general

El modelo de regresión múltiple general considera a 759 estudiantes que no tomaron la tutoría y a 241 que sí lo hicieron. Para determinar los límites inferior y superior en el punto de corte se calculan modelos de regresión para cada uno de los grupos, estableciendo la ecuación de regresión de las dos líneas de tendencia observadas.

Ecuación recta de regresión del grupo control:

$$\text{Examen final} = 59.4 + 0.51 (\text{examen de entrada centrado}) + \beta_2 \text{tutoría} + \varepsilon$$

El valor 59.4 indica el punto de corte de la línea de tendencia β_0 de quienes no participaron en el programa y se identifica como el límite inferior de la distancia vertical que denota el impacto del programa.

Ecuación recta de regresión del grupo de tratamiento:

$$\text{Examen final} = 70.20 + 0.509 (\text{examen de entrada centrado}) + \beta_2 \text{tutoría} + \varepsilon$$

El impacto del programa se mide sobre la línea vertical definida por el umbral y se determina como la diferencia entre los dos puntos de las líneas de tendencia de los grupos bajo estudio. Es una medida vertical cuyo límite superior está determinado por el punto máximo al-

canzado, β_0 , de la ecuación de regresión calculada con los puntajes obtenidos por los estudiantes que tomaron la tutoría, el cual es: 70.2. Por tanto, el impacto es la diferencia entre estos dos puntos ($70.2 - 59.4 = 10.8$). Significa que los alumnos que recibieron la tutoría mejoraron, en promedio, su puntaje de examen final en 10.8.

Ahora bien, qué sucede cuando la evaluación se reduce a los estudiantes que están muy cerca del punto de corte. Este es un interrogante importante, pues la diferencia entre quienes están ligeramente por encima o por debajo, son muy similares y lo único que los diferencia es el tratamiento recibido (tutoría). No debe olvidarse que si la diferencia umbral es más reducida, la similitud de las observaciones es mayor. Estableciendo anchos de banda, entonces, la comparación entre dos grupos es más adecuada y el impacto puede apreciarse de mejor manera. Esta es una manera de establecer el contrafactual para hacer comparaciones juiciosas.

Análisis de regresión local

En el contexto del análisis de la regresión discontinua, se introduce el análisis de regresión lineal local como un procedimiento que permite estimar una regresión en la vecindad del punto del umbral, calculando los parámetros de la línea de tendencia en cada grupo de estudio (Jacob y Zhu, 2016). Este procedimiento consiste en determinar un ancho de banda (*bin or bandwidth*), a cada lado del punto de corte o umbral (izquierda-derecha). De esta manera, el analista reduce la influencia de los valores extremos a ambos lados del límite establecido y concentra su esfuerzo en estudiar el impacto en una zona más homogénea de puntajes, cuya diferencia en puntajes solamente puede ser aducida al tratamiento aplicado. Así, por ejemplo, para el análisis propuesto establece anchos de banda alrededor del límite establecido para recibir la tutoría (70 puntos). Entonces, para un ancho de banda de 10, los puntajes límite serán 60 a la izquierda y 80 a la derecha; para un ancho de banda de 5, serán de 65 y 75, respectivamente. Considerando el examen centrado, el umbral es cero y el ancho de banda 10 está comprendido entre -10 y $+10$, mientras que un ancho de banda 5 estará en el rango -5 y $+5$ puntos. Obviamente, este análisis considera un número menor de observaciones y calcula de nuevo los puntos de corte en el umbral de las dos líneas de tendencia para establecer la medida del impacto.

Ancho de banda 10

Como se había anticipado, este análisis de regresión local comprende los puntajes obtenidos por los estudiantes en un rango entre 60 y 80 puntos de los exámenes de los puntajes iniciales (entre -10 y $+10$ del examen centrado), una franja más reducida para estimar el impacto. El número de observaciones para este análisis se reduce de manera significativa y solo se consideran 149 alumnos del grupo de tratamiento y 255 del grupo de control. Con las observaciones seleccionadas, de nuevo se calculan modelos de regresión lineal para conocer las ecua-

ciones de las respectivas rectas. Como se sabe, cuando la variable x toma el valor cero determina el punto de corte de la recta de regresión en y , en este caso el umbral. Este punto de corte está determinado por el parámetro β_0 . Para un ancho de banda 10, la recta del grupo de control corta el eje vertical en 60,202, determinando así el límite inferior de la distancia vertical del impacto. En el grupo de tratamiento, la línea de tendencia presenta un corte en 69,424, determinando el límite superior del impacto. En otras palabras, el impacto está determinado por la diferencia entre los límites superior e inferior de esta distancia vertical sobre el umbral y es de 9.22 (tablas 10.4 y 10.5).

TABLA 10.4. MODELO DE REGRESIÓN GRUPO DE CONTROL
PARA ANCHO DE BANDA 10

Coeficientes ^a						
Modelo		Coeficientes no estandarizados		Coef. estand.	t	Sig.
		B	Desv. error	Beta		
I	(Constante)	60,202	893		67,435	
	examen centrado	,420	144	181	2,920	4

^a Variable dependiente: Examen final.

Fuente: elaboración propia a partir de SPSS.

TABLA 10.5. MODELO DE REGRESIÓN GRUPO DE TRATAMIENTO
PARA ANCHO DE BANDA 10

Coeficientes ^a						
Modelo		Coeficientes no estandarizados		Coef. estand.	t	Sig.
		B	Desv. error	Beta		
I	(Constante)	69,424	975		71,202	
	examen centrado	,336	189	145	1,780	77

^a Variable dependiente: Examen final.

Fuente: elaboración propia a partir de SPSS.

Ancho de banda 5

Un ejercicio más juicioso y contar así con puntos de comparación para la estimación del impacto, presenta un análisis de regresión con las observaciones comprendidas en un ancho de banda 5. Este considera las observaciones a cinco puntos por debajo del umbral y cinco puntos por encima. De nuevo se establecen modelos de regresión local para encontrar la ecuación de la recta de regresión y el punto de corte con la línea de corte. Para una banda de 5, el límite inferior de la tendencia del grupo de control (cuando x toma el valor cero) se fija en 60.1. Para el grupo de tratamiento, su línea de tendencia presenta la intersección con el punto de corte en 69.37, determinando el límite superior del impacto. Para un ancho de banda 5, el impacto se estima en 9.25, la diferencia entre los límites superior (69.39) e inferior (60.14) (tablas 10.6 y 10.7).

TABLA 10.6. MODELO DE REGRESIÓN LOCAL GRUPO DE CONTROL
PARA ANCHO DE BANDA 5

Coeficientes ^a						
Modelo		Coeficientes no estandarizados		Coef. estand.	t	Sig.
		B	Desv. error	Beta		
I	(Constante)	60,141	1,422		42,295	
	examen centrado	583	472	121	1,236	219

^a Variable dependiente: Examen final.

Fuente: elaboración propia a partir de SPSS.

TABLA 10.7. MODELO DE REGRESIÓN LOCAL GRUPO DE TRATAMIENTO
PARA ANCHO DE BANDA 5

Coeficientes ^a						
Modelo		Coeficientes no estandarizados		Coef. estand.	t	Sig.
		B	Desv. error	Beta		
I	(Constante)	69,397	1,320		52,564	
	examen centrado	219	491	49	447	656

^a Variable dependiente: Examen final.

Fuente: elaboración propia a partir de SPSS.

Paso 6. Establecer las comparaciones del caso.

Las distintas evaluaciones del impacto del programa de tutoría se compilan en la tabla 10.8, permitiendo una visión y análisis global.

TABLA 10.8. COMPARACIÓN RESULTADOS DE EVALUACIÓN DEL PROGRAMA DE TUTORÍA

Modelo	Límite superior	Límite inferior	Impacto	N
General	70.20	59.40	10.8	1,000
Ancho de banda (10)	69.42	60.20	9.22	404
Ancho de banda (5)	69.39	60.14	9.25	191
Promedios	69.67	59.91	9.76	

Fuente: elaboración propia.

Con las distintas versiones de la medición del impacto se observó un rango entre 9.22 y 10.8 puntos promedio que los estudiantes que recibieron la tutoría podrían mejorar en sus pruebas finales. Estas medidas evidencian la efectividad del programa y evalúan su impacto, considerando distintas observaciones alrededor del punto de corte (umbral). La incorporación del ancho de banda con distintas medidas permite, por un lado, depurar el efecto de los puntos distantes del umbral y, por otro lado, precisar la medida del impacto del programa para las observaciones que se encuentran arriba/abajo de dicho umbral. En resumen, la regresión discontinua local proporciona precisión a la medición del impacto, dándole al analista mayor certeza sobre su verdadera magnitud.

10.6 SÍNTESIS CAPITULAR

El presente capítulo sobre la evaluación de impacto deja varias reflexiones para consideración de los evaluadores y los *stakeholders* de las políticas. Lo primero es tomar conciencia sobre los retos que entraña la evaluación del impacto; sin duda, un ejercicio desafiante que exige el mayor rigor de los analistas. Segundo, como en cualquier ejercicio estratégico, es requisito indispensable formular, desde el diseño del programa, los criterios y condiciones para evaluar el impacto, garantizando la disposición de la información necesaria y, con ello, propiciar una cultura de la evaluación. Por último, insistir en la necesidad de generar una masa crítica de analistas para que la evaluación sea un ejercicio regular, riguroso y útil para la toma de decisiones, es decir, que impriman cierto rigor científico.

REFERENCIAS

- Abad, H. (2009). *Traiciones de la memoria*. Alfaguara.
- Baker, J. L. (2000). *Evaluating the Impact of Development Projects on Poverty. A Handbook for Practitioners*. The World Bank.
- Cattaneo, M. D.; Idrobo, N. y Titiunik, R. (2019). *A Practical Introduction to Regression Discontinuity Designs: Foundations*. Cambridge Elements: Quantitative and Computational Methods for Social Science, Cambridge University Press. <https://arxiv.org/abs/1911.09511>
- Gertler, P. J.; Martínez, S.; Premand, P.; Rawlings, L. B. y Vermeersch, C. M. J. (2016). *Impact Evaluation in Practice*. Washington: International Bank for Reconstruction and Development/The World Bank. <https://doi.org/10.1596/978-1-4648-0779-4>
- Heiss, A. (2020). Program Evaluation. Andrew Young School of Policy Studies, Georgia State University, <https://evalf20.classes.andrewheiss.com/example/rdd/#program-background>
- Jacob, R. y Zhu, P. (2012). *A Practical Guide to Regression Discontinuity*, Michigan University. www.mdrc.org
- Leeuw, F. y Vaessen, J. (2009). *Impact Evaluations and Development. Nonie Guidance on Impact Evaluation*. World Bank Group.
- Pelegrín, J. (2010). “La historia alternativa como herramienta didáctica: una revisión historiográfica”, *Proyecto CLIO*, 36. <http://clio.rediris.es>
- Swiss Agency for Development and Cooperation, SDC (s/f). *What are Impact Evaluations?* In collaboration with Nadel Ethz, Center for Development and Cooperation.

