

La evaluación de políticas públicas: enfoques teóricos y realidades en nueve países desarrollados

Juan Pablo Guerrero Amparán

Al señalar los enormes poderes discrecionales del ejecutivo en México, Carlos Fuentes lamenta que palabras usuales del derecho público anglosajón como *accountability* y *checks and balances* sean de tan difícil traducción al idioma de los mexicanos. No obstante, el escritor traduce dichos conceptos del siguiente modo: “rendimiento de cuentas” y “límites y contrapesos”.¹ Éstos refieren a la posibilidad de tener el derecho y la capacidad efectiva de exigir a una autoridad pública que explique sus acciones, lo cual no sólo requiere de marcos legales que lo hagan posible sino de instituciones independientes con el poder suficiente para obligarla a hacerlo.

El hecho de que los integrantes de un gobierno y de la administración sean responsables de sus actividades ante el público, sean punibles y rindan cuentas a los ciudadanos o cuerpos de poder autónomos, es una de las características fundamentales de los sistemas políticos democráticos actuales. En Francia, Canadá o Estados Unidos, cuando se comprueba que un jefe de gobierno transgredió la ley, otras instancias públicas son capaces de hacerle pagar un costo legal y político por ello. En las líneas siguientes nos vamos a referir principalmente a la noción del control sobre los ejecutantes de programas gubernamentales a través de la evaluación.²

Juan Pablo Guerrero Amparán es investigador del CIDE y candidato a doctor por el Instituto de Estudios Políticos de París.

¹ Desde luego, Carlos Fuentes lamenta también que *accountability* y *checks and balances* no se traduzcan en la práctica política mexicana. Véase el diario mexicano *La Jornada*, 11 de enero de 1995.

² Queremos decir evaluación de resultados. Existen diversos términos que denominan al

En décadas recientes, las instituciones públicas en varios países, dependientes del ejecutivo pero también de otras ramas del Estado, han recurrido a los métodos de las ciencias sociales para supervisar el funcionamiento de los programas públicos. Ello ha impulsado el desarrollo de distintas técnicas y enfoques para la evaluación. En términos generales, escriben Yves Mény y Jean-Claude Thoenig, los especialistas del estudio de la actuación del Estado han asumido dos actitudes frente a la evaluación. Una de ellas es determinar si existen, dentro de un sistema político, los organismos o instancias que puedan actuar efectivamente como órganos de evaluación.³ La otra consiste en “construir los conceptos y los métodos que permiten describir, explicar y juzgar en forma científica los resultados de la acción gubernamental”.⁴

En este artículo nos ocuparemos de ambas opciones: en la primera parte repasaremos los elementos clave de la discusión sobre los procedimientos para hacer una evaluación de programas estatales desde la perspectiva de la disciplina del análisis de políticas públicas. En la segunda, nos aplicaremos a la revisión de la realidad de los procesos de evaluación en nueve sistemas políticos: Estados Unidos, Alemania, Canadá, Reino Unido, Dinamarca, Suiza, Holanda, Noruega y Francia.⁵ Vamos a describir y a comparar las estructuras institucionales de evaluación y control de sus respectivos aparatos administrativos estatales.

genérico *evaluación*: “análisis de política”, “evaluación de políticas”, “valuación de programas”, “estudio de políticas”, “evaluación *ex ante* o *a priori*”, “*ex post* o *a posteriori*”, “monitoreo de políticas”, “auditoría de eficacia”, “pronóstico de políticas”, etc. Aquí haremos la distinción entre *evaluación de resultados*, que estudia en retrospectiva las políticas y programas gubernamentales, o las estudia paso a paso mientras se aplican, y *análisis de políticas (policy analysis)*, que busca pronosticar los posibles efectos de una acción del gobierno para asesorar o advertir al que decide en el ámbito público. Una discusión clara al respecto está en Myriam Cardozo Brum, “La evaluación de las políticas públicas: problemas, metodologías, aportes y limitaciones”, *Revista Mexicana de Administración Pública*, México, 1993, pp. 167-197.

³ Desde instituciones más o menos independientes del ejecutivo, encargadas de llevar a cabo las evaluaciones, hasta organizaciones civiles no gubernamentales, pasando por empresas consultoras que se contratan *ex profeso* y órganos de opinión pública, como la prensa escrita. Lo anterior puede consultarse en Y. Mény y J. C. Thoenig, *Politiques publiques*, Paris, Presses Universitaires de France, 1989, pp. 287-325.

⁴ Mény y Thoenig, *op. cit.*, p. 288.

⁵ Nos basaremos principalmente en el estudio al respecto al que convocó Ray C. Rist y que se publicó en *Program Evaluation and the Management of Government. Patterns & Prospects across Eight Nations*, New Brunswick, N.J., Transaction Publishers, 1990, 184 p. Para el caso de Francia, consideraremos a Jean-Pierre Nioche y Robert Poinard, “Public Policy Evaluation in France”, *Journal of Policy Analysis and Management*, vol. 5, núm. 1, 1985, pp. 58-72.

Primera parte. Los enfoques de evaluación en el análisis de políticas públicas

La evaluación forma parte de la vida de cualquier organización, aunque suceda de manera tácita, y consiste en que sus integrantes identifiquen y midan los efectos de las acciones individuales y del conjunto. Si nos referimos a la acción pública de un gobierno, no solamente sus integrantes sino particularmente los sectores sociales que se ven afectados por ella, se preguntan acerca de sus resultados, su justificación y su efectividad. Cualquier ciudadano, al observar la actividad gubernamental, puede preguntarse: ¿se justifica el pago de mis impuestos? ¿Cumple el gobierno con las expectativas que promovió? ¿Ofrece con eficiencia y justicia los servicios públicos? ¿Usa con honestidad y transparencia los recursos con los que cuenta? ¿Es necesaria tal o cual intervención del Estado? ¿Resuelve o crea problemas?

No hace mucho, una referencia frecuente para medir la efectividad del gobierno en México era la cantidad de infraestructura creada. Si un gobierno construía escuelas, caminos, puentes y obra hidráulica, la comunidad sentía que el gobierno era suficientemente bueno. Pero a partir de los años ochenta, cuando el Estado mexicano sufrió una escasez de recursos, tanto los integrantes de la administración como algunos sectores de la opinión pública empezaron a mostrarse más interesados por los problemas de equidad de la acción pública, así como por la eficiencia, la rentabilidad y la calidad de los servicios.⁶ Un poco más tarde, el público comenzó a preocuparse no sólo por la cantidad y la calidad de las acciones gubernamentales y las obras públicas, sino también por su fundamentación legal, su justificación moral y el grado de responsabilidad de los funcionarios que deciden y llevan a cabo los programas en cuestión.

Efectivamente, muchos pueden ser los criterios para evaluar una política pública.⁷ Si revisamos los orígenes del concepto de evaluación vemos que tiene sus raíces etimológicas en la palabra “valor”. Pero

⁶ Un estudio reciente sobre la prestación de servicios públicos muestra que el grado de calidad y el problema de la cobertura son serias preocupaciones entre los funcionarios municipales del Estado de México, por poner un ejemplo. Véase Rodolfo García del Castillo, *Cobertura y calidad de los servicios municipales. Estudio sobre el Estado de México* (Documento de Trabajo 12, AP), México, CIDE, 1993, 93 p.

⁷ Existen infinidad de definiciones para el término *políticas públicas*. Casi todas tienen un denominador común, que Jarol B. Manheim y Richard C. Rich recogen así: “son las actividades o inactividades regulares y tipificadas de los funcionarios del Estado”. Véase *Análisis político empírico; métodos de investigación en ciencia política*, Madrid, Alianza, 1988, p. 428.

sabemos, asimismo, que el valor de un objeto, de una idea, de una palabra o de una acción siempre es relativo a la referencia o escala que se usa para apreciarla.⁸ Para algunos investigadores de las acciones del gobierno, la evaluación es una fase de un proceso amplio que constituye la intervención de la autoridad pública.⁹ En general, no se limita a describir simplemente los sucesos ligados a la aplicación de una política determinada o a su calificación de acuerdo con determinadas normas o criterios contables y presupuestales; también pretende explicar por qué tuvo el efecto logrado y examinar si el problema que originó la actuación gubernamental fue finalmente resuelto y si se alcanzaron sus objetivos (los explícitos y los implícitos). Más aún, el evaluador quiere comprender las posibles transformaciones en las relaciones entre las instituciones ejecutantes y los sectores sociales involucrados, los efectos inesperados de la actuación pública, y las transformaciones internas que la propia autoridad experimentó al poner en práctica su programa.

Pretender evaluar de manera objetiva —y hasta científica— los resultados de la acción gubernamental ha implicado para el analista la construcción de conceptos y métodos muy diversos. Para apreciar los efectos políticos, económicos y sociales generados por una intervención gubernamental, en un área específica de la sociedad y en un lugar y tiempo determinados, el evaluador escoge un procedimiento, un método de observación y una forma de razonar específicos. En realidad

⁸ A su vez, la determinación de la referencia, escala o índice de evaluación depende de varias condiciones objetivas y, sobre todo, de consideraciones subjetivas: la situación particular (emocional y material) de quien califica, el contexto general en el que se encuentra, las creencias individuales, la moral predominante, el marco institucional y jurídico, etcétera.

⁹ Varios autores han desarrollado la noción de *proceso de una política pública* a partir de las preocupaciones por su elaboración, su orientación hacia los problemas sociales y su aplicación dentro de un contexto social determinado. Ellos parten de cierta personificación del Estado y asumen que su acción sigue, a grandes rasgos, el proceso mental de la acción humana, desde que el actor percibe un motivo y decide que debe actuar, hasta que finaliza su acción (después de pensar en alternativas, tomar una determinación, llevarla a cabo, observar las consecuencias y dar por terminada su actuación). Entre los distintos modelos —que no son sino instrumentos conceptuales para facilitar la interpretación de la realidad del Estado en acción— que se han propuesto para dar cuenta de la vida de las políticas públicas, casi todos, al reflejar también un ideal, incluyen una fase de evaluación. Harold D. Lasswell se refirió a varias fases (apreciación, promoción, prescripción, innovación, aplicación, evaluación, terminación, valoración) desde 1956 en *The Decision Process*; R. May y A. Wildavsky identifican cinco fases (construcción de la agenda, análisis, ejecución, evaluación, terminación); Gary Brewer y Peter deLeon hablan de seis etapas (iniciación, estimación, selección, instrumentación, evaluación, terminación); en *The Foundations of Policy Analysis*, California, Brooks/Cole, 1983, 338 p. P. deLeon también se refiere a ellos en su artículo en *Foro Internacional*, 131, p. 97. Para todos, la evaluación es una fase integral de la vida de la política pública.

muchos son los enfoques que se han desarrollado en torno a este tema. Según Mény y Thoenig, las diferentes estrategias para evaluar se pueden clasificar en cuatro enfoques generales: el descriptivo, el clínico, el normativo y el experimental. Nosotros hemos agregado la evaluación de tipo *pluralista*, que combina algunos de los criterios de los otros enfoques y ofrece al evaluador un papel activo a lo largo de la instrumentación de la política.¹⁰

Antes de describir los rasgos distintivos de cada uno de los enfoques es necesario hacer una revisión rápida de cinco conceptos que son básicos en la realización de una evaluación. El primero es el de los *objetivos* generales de la política, los cuales se refieren a las metas, fines, circunstancias o situaciones que se buscan, ya sea porque se consideren deseables por sí mismos o porque resulten prácticos. Junto a los objetivos del programa gubernamental toda evaluación también tiene sus *objetivos*, que pueden ser los mismos del programa gubernamental o los que el evaluador seleccione. Es decir, pueden ser dependientes de las metas del programa gubernamental cuando estudian si éstas en efecto se cumplieron, o bien relativamente independientes de los objetivos oficiales cuando estudian los resultados desde una óptica externa y con criterios autónomos. Aquí nos ocuparemos de los objetivos de la evaluación, haciendo énfasis en la compleja relación que se establece con las metas oficiales.¹¹

El hecho de que los objetivos se deriven de la necesidad o de la virtud, depende del responsable de la decisión; pero ninguna evaluación puede pasar por alto las motivaciones de la actuación pública. Para evaluar los objetivos y los logros, las metas y las consecuencias de la acción se establecen en segundo lugar los *criterios*, que son las “reglas para conocer la verdad”: referencias o escalas que sirven para graduar, interpretar y calificar la realidad. Los criterios de la evaluación se llevan a efecto —se operacionalizan— a través de métodos de análisis, observación, registro y medición particulares. En efecto, existen una infinidad de procedimientos, medios e instrumentos que per-

¹⁰ La evaluación puede ser *interna*, cuando la elaboran instancias que son dependientes de la organización que aplica la política, o *externa*, cuando la realiza un organismo que no está integrado y es relativamente autónomo respecto del cuerpo público responsable de la política a evaluar. Asimismo, puede ser considerada como una fase de la política pública separada del resto, aunque en realidad las fases no sean fácilmente distinguibles unas de otras, pues se confunden y empalman con frecuencia.

¹¹ Desde su perspectiva de estudio Myriam Cardozo identifica *seis objetivos organizacionales* de las evaluaciones: 1) microeconómicos y financieros; 2) macroeconómicos; 3) tecnológicos; 4) sociales; 5) objetivos de equidad; y 6) políticos. Véase M. Cardozo, *op. cit.*, p. 169.

miten poner en práctica la evaluación: son las *técnicas* o *métodos* de evaluación, que se derivan, generalmente en forma coherente, de los criterios.¹²

El objeto de la evaluación es proporcionar una graduación (o puntuación), interpretación o calificación a la política pública; ello se traduce en *valores*, que es el cuarto concepto.¹³ Es con base en los criterios que pueden determinarse las características, medirse los méritos y calcularse los desaciertos de un programa de actuación gubernamental: el valor puede remitir a la importancia, la calidad, la sustancia, la cantidad, la estimación, la utilidad, la justicia, la eficacia, etc., de la política pública.¹⁴

Finalmente, cuando varios individuos aceptan una determinada jerarquización de valores (lo bueno y lo malo, lo superior y lo inferior, lo que cuesta y lo que no, etc.) y cuando, a través de la costumbre, por ejemplo, los valores se implantan en la moral convencional, surgen las *normas* (aunque en una organización social, pueden ser definidas impositivamente, *por ley*). Las normas son preceptos de conducta que gobiernan y dirigen el comportamiento al asegurar la regulación de la vida de los individuos y los grupos dentro de la sociedad.

Objetivos, criterios, técnicas, valor y normas son los conceptos de partida para la elaboración de las evaluaciones. A continuación, men-

¹² Para la evaluación de los efectos de una política pública, M. Cardozo distingue sólo tres: análisis de excedente de productividad, balances sociales y estudios *cuasiexperimentales*. Pero la autora reconoce que varias de las técnicas que identifica para el análisis de políticas (evaluación *ex ante*) también pueden considerarse para la evaluación de resultados, aunque algunas de ellas pierdan su ideal "optimizador". El énfasis para nosotros es que para cada criterio existe una técnica que le corresponde en forma más o menos adecuada. Las técnicas aludidas en el texto de Cardozo (*op. cit.*, p.184) son: árboles de decisión, análisis financiero, análisis costo-beneficio, análisis costo-eficacia y teoría de juegos.

¹³ En términos generales, son tres los métodos para establecer los criterios de valoración: 1) El nominal, que consiste en dar un nombre a un objeto, grupo, sector, proceso, situación, con el fin de clasificarlo. El nombre debe ser un calificativo: bueno-malo, eficaz-inerte, falso-verdadero, para lo cual basta aclarar su significado, con base en la extensión buscada (número de cosas que señala, a las que se refiere) y a su connotación (su significado, el número de propiedades e ideas a las que hace referencia); con dichas especificaciones se le puede clasificar en una escala de valores determinada. 2) El método ordinal, cuando se expresa una sucesión, alguna disposición ordenada de las partes de un todo, una relación con base en características de calidad o de cantidad (primero, segundo..., más, menos, etc.) 3) Y el cardinal, que atribuye un adjetivo numérico que expresa la cantidad o el grado de cosas de que se trata, de acuerdo con características cualitativas o cuantitativas, con el fin de situar en una escala (con base en la escala de 10, por ejemplo).

¹⁴ A partir de ideales (prototipos o modelos de perfección), toda sociedad se organiza con base en un sistema de valores que debe ser relativamente congruente con la ideología predominante, las creencias y los mitos de los grupos sociales. Al mismo tiempo, los valores deben conjugarse lógicamente con la moral práctica y las normas, mismas que regulan la organización y el funcionamiento de la vida política, social y económica.

cionaremos los rasgos distintivos de cada uno de los enfoques generales que se ocupan de esta cuestión.

El enfoque descriptivo elabora una lista precisa y exhaustiva de los efectos de la acción pública. Su propósito es registrar los cambios generados por la política pública, tal como lo haría una cámara fotográfica al plasmar la realidad antes y después de la aplicación del programa público. Tras el periodo establecido para llevar a cabo la evaluación (por ejemplo, el que transcurre entre el inicio de la aplicación de la política y el momento escogido para evaluarlo), se establece un inventario de cada una de las variaciones que registraron en el sector social en el que se aplicó dicha política.

Generalmente, los enfoques descriptivos se dedican al estudio del entorno social que la política pública buscó modificar (más que al aparato institucional que la puso en práctica), y se pretenden neutrales: registran cada cambio pero no miden el grado de variación, ni juzgan su ocasional efectividad (ni condenan su eventual inoperancia). Este enfoque de evaluación estudia principalmente los efectos de la política en cuestión, es decir, las perturbaciones que ocurren tras la intervención de la autoridad.

También pueden identificarse los efectos de una política preventiva, que busca precisamente "modificar" el curso esperado de los sucesos sociales —curso "natural" o "lógico" de las cosas—. Toda autoridad pública tiene una cierta idea de lo que puede ocurrir, de acuerdo con su experiencia, sus conocimientos previos y la información con que cuenta sobre el contexto en que se encuentra y que desea modificar o preservar. Con esos puntos de partida, la autoridad puede verse inclinada a anticiparse o a prevenir ciertos sucesos. Un ejemplo sencillo es el caso de un incremento de salarios a trabajadores sindicalizados, o el encarcelamiento de su líder por un fraude fiscal a comprobar posteriormente, para evitar una huelga en el sector público: los efectos de la actuación gubernamental pueden ser medidos concretamente, pero también en forma abstracta o simbólica (opiniones de la gente, su nivel de satisfacción al pensar en la política pública estudiada, las percepciones y las creencias, etc.).¹⁵ Por otro lado, el enfoque descriptivo también se ha interesado en registrar los cambios o la conservación que una política genera en el medio ambiente y en el territorio (efecto de

¹⁵ Véase Mény y Thoenig, *op. cit.*, p. 292.

la construcción de un desarrollo turístico, creación de parques nacionales, etcétera).

Las técnicas de observación para la evaluación descriptiva son muy variadas. Las opciones a las que el investigador puede recurrir se parecen bastante a las que se encuentran en los manuales de base de las ciencias sociales en general: elaboración de muestreos, investigación por encuestas y circulación de cuestionarios, análisis estadísticos, observación participante o directa, análisis de contenido (informes gubernamentales), construcción de escalas o de modelos matemáticos, etc.¹⁶ Pero en este enfoque, lo importante es observar las transformaciones o desviaciones del "curso normal" de las cosas que son atribuibles a la política pública en cuestión, es decir, lo que cambió gracias a la actuación del gobierno. Es muy común que los evaluadores no expresen de manera explícita los criterios que los guían, pero tienen que exhibir de alguna forma las técnicas usadas. Esto resulta muy útil para quien realiza un análisis de la evaluación, pues generalmente las técnicas traducen en varios sentidos los criterios subyacentes de los evaluadores.

Cabe anotar que al describir los efectos es inevitable partir de la observación de lo que se esperaba modificar; por ello en la descripción final es frecuente que la evaluación descriptiva subraye los efectos inesperados. Cuando se han observado los efectos de uno de los programas de lucha contra la contaminación ambiental en la ciudad de México, por ejemplo el programa "Hoy no circula" (que consiste en que cada automóvil tiene prohibido circular uno de los cinco días hábiles semanales), se ha visto que al principio, más allá de su efectividad en cuanto al decremento de la contaminación por una supuesta reducción del tráfico de vehículos, la demanda en el mercado de automóviles usados resultó relativamente favorecida, pues un número considerable de propietarios de autos de ingresos medios y altos optó por la adquisición de uno adicional, en lugar de organizar turnos de viajes entre los vecinos o usar con más frecuencia el transporte público.¹⁷

Un ejemplo de la evaluación descriptiva es cierto tipo de estudio

¹⁶ Hay un sinnúmero de títulos que tratan sobre estas técnicas. Al respecto sugerimos la revisión, por recomendación del investigador del CIDE, Gerardo Rivera Navarro, de Jarol B. Manheim y Richard C. Rich, *op. cit.*

¹⁷ Vale decir que, respecto al programa "Hoy no circula", no hay evaluaciones que hasta el momento se hayan hecho públicas, aunque sus efectos inesperados han sido registrados en numerosas ocasiones. Mény y Thoenig (*op. cit.*, p. 295) identifican tres causas principales para explicar las consecuencias inesperadas de la intervención del Estado: 1) La autoridad que diseñó la política no entendió correctamente las causas del problema que quiso atacar. 2) Los problemas

de caso de un programa público en determinado sector social, en un periodo establecido y en un lugar geográfico específico. La observación es particular, sin que exista la posibilidad de establecer criterios comparativos con base en la observación de un grupo similar en el que no se hubiera aplicado dicha política. En el *estudio de caso*, el evaluador tiene solamente la información correspondiente al grupo afectado por la política evaluada: debe conformarse con observar los efectos de la política y comparar la situación con la que prevalecía antes de la aplicación de la política (comparación temporal) o bien, en el mejor de los casos, hacer un seguimiento progresivo a lo largo de la aplicación de la política, con evaluaciones constantes (análisis de series temporales). El establecimiento de las relaciones de causalidad se dificulta al no haber un grupo de referencia para establecer un análisis comparado (en realidad no se sabe qué hubiera pasado si no se hubiera aplicado esa política). Por ejemplo, siempre existe el riesgo de que lo que se interpreta como un efecto sea en realidad la influencia de factores exógenos a la política aplicada.

Sin embargo, este enfoque no busca en realidad establecer ese tipo de asociaciones entre causas supuestas y consecuencias concretas. Existen técnicas analíticas de observación suficientemente aceptadas que permiten distinguir las modalidades de aplicación de la política pública evaluada o se proponen interpretaciones sobre sus efectos, pero ello forma parte de otros enfoques de evaluación, como los que veremos a continuación. En la evaluación descriptiva debe mantenerse una gran modestia en cuanto a su interpretación teórica, así como evitarse cualquier intento de generalización.

El enfoque clínico, como los enfoques que presentaremos posteriormente, busca explicar los efectos: parte de los objetivos declarados de la política pública y, al observar los resultados obtenidos, intenta explicar por qué algunos de los objetivos no se lograron y por qué algunas consecuencias resultaron inesperadas. Los criterios de evaluación son *endógenos*, definidos a partir de las metas propuestas por los diseñadores de la política en cuestión. Este enfoque, como el descriptivo, registra

evolucieron y por lo tanto el entorno que se pretendía solucionar cambió (generalmente sucede cuando la acción gubernamental llega demasiado tarde; también sucede cuando otras acciones públicas o factores externos provocan efectos en el sector considerado). 3) Los cambios introducidos por la autoridad generaron consecuencias colectivas incontrolables o gran fluidez simultánea en varios sectores sociales. Véase en este tercer punto a Thomas Schelling, *La tyrannie des petites décisions*, París, Presses Universitaires de France, 1980.

también las variaciones provocadas por la intervención de la autoridad pública en un determinado periodo, pero establece su análisis deliberadamente a partir de lo que se buscaba modificar con la política, por lo que hace énfasis en los criterios iniciales de la misma. Identifica las diferencias entre el objetivo verdadero (muchas veces tácito) y el buscado (explícito u "oficial") por el funcionario que efectúa la política respecto al efecto realmente alcanzado, y trata de explicar, a través del análisis de las condiciones previas a la aplicación de la política y de la consideración de los problemas a lo largo de su instrumentación, qué fue lo que falló. Generalmente, como el descriptivo, el enfoque clínico es una evaluación posterior (*ex post*) a la aplicación de la política y se centra en sus efectos.

Este enfoque no se pretende crítico, pues desde un principio hace suyos los objetivos de la política pública en cuestión. Al contrario, más allá de consideraciones morales, su pretensión es ser útil: sirve a los objetivos del funcionario responsable de la política pública, pues busca, a través de una interpretación práctica, determinar las causas de las variaciones entre objetivos y resultados a fin de evitarlas en el futuro. Este enfoque en realidad califica, con valores que son congruentes con los criterios internos de la entidad pública actuante, el cumplimiento de los objetivos: va y viene entre sus fines últimos y sus resultados concretos, de manera que la evaluación se convierte en la medición que le corresponde a la política de la variación de una situación, a diferencia de la que le corresponde a los factores exteriores a esa política específica. Ahí residen, por lo demás, sus limitaciones, pues por lo general resulta muy difícil discernir los cambios que se deben a la política y los que resultan de la modificación de parámetros exteriores, o de otra política pública que actuó simultáneamente e influyó el resultado final en forma indirecta.

Existen tres técnicas más o menos sistemáticas que se utilizan frecuentemente para este tipo de evaluación. La encuesta estadística, que establece pruebas sobre sectores representativos del grupo social afectado (lo cual supone que se tiene un buen conocimiento estadístico de ese grupo) y permite establecer una tipología de comportamientos frente a la política. Incluso, en ciertas condiciones, esta técnica permite establecer relaciones de causalidad, por lo cual algunos autores la identifican como *cuasiexperimental*.¹⁸ La segunda técnica ya ha sido men-

cionada: la evaluación antes y después de la aplicación de la política. Se le denomina *evaluación sumaria* (*summative evaluation*), y consiste en un análisis de las consecuencias generales de la política en cuestión, ya que pone el énfasis en los efectos finales, después de considerar la situación que prevalecía con anterioridad. Finalmente, existe el seguimiento progresivo de la política aplicada, lo cual puede lograrse cuando se instala un grupo de observación permanente (como los de las áreas de la formación¹⁹) o en el caso de políticas públicas que mantienen constantes sus objetivos pero reajustan progresivamente sus medios, en vista de los resultados sucesivos de la evaluación. Este tipo de evaluación se conoce como *formativa* (*formative evaluation*), pues permite establecer observaciones y recomendaciones tácticas a los responsables de la política (los que decidieron ponerla en práctica) con el fin de corregir desviaciones y mejorar el desempeño del programa. Las evaluaciones sumarias hacen énfasis en los efectos reales del programa gubernamental evaluado, mientras que las formativas cuidan más de cerca el cumplimiento de los efectos buscados.²⁰

No siempre es fácil establecer una evaluación clínica, pues sus supuestos entienden que los objetivos de una política pública están claramente definidos (lo cual no siempre es cierto, pues generalmente existen los objetivos declarados oficialmente y los objetivos reales, pero escondidos del grupo de autoridades), son relativamente coherentes y en verdad los buscan sus impulsores. Sin embargo, no es raro que una misma política pública exhiba dos fines que sean más o menos excluyentes entre sí, o que sean francamente opuestos. En otras ocasiones, los medios usados resultan contradictorios respecto del objetivo buscado. Por ejemplo, dos objetivos con frecuencia opuestos en una política económica de fomento al crecimiento en México con estabilidad de pre-

afectados suficientemente heterogéneos como pilotos de prueba, para que la estimación econométrica pueda aislar los efectos de la política, al sacar provecho de las diferencias entre las observaciones de cada corte transversal. Los pioneros en desarrollar las técnicas estadísticas para interpretar los efectos de la intervención gubernamental en análisis de series de tiempo (mediciones sucesivas de las variables para determinar tendencias) fueron G.E. Box y Gwilyn Jenkins, en 1970 (*Time Series Analysis*, San Francisco, Holden-Day).

¹⁹ En la Secretaría de Educación Pública, por ejemplo, existe la Comisión Nacional Mixta de Escalafón, que controla a los inspectores y encargados de evaluar al personal docente.

²⁰ A la evaluación formativa, Eric Monnier la llama *endoformativa*: sus objetivos son la concepción y mejora de la innovación y los destinatarios son los protagonistas de la innovación. A la evaluación sumaria, Monnier la denomina *recapitulativa*, cuyo objetivo es validar o garantizar la innovación, y los destinatarios son los responsables nacionales, colegas expertos, profesionales y usuarios potenciales. Véase "Objetivos y destinatarios de las evaluaciones", *Documentación Administrativa*, núm. 224-225, octubre de 1990-marzo de 1991, p. 148.

¹⁸ Implica, por ejemplo, una modelización microeconómica, con un conjunto de individuos

cios (por lo menos, muy difícilmente compatibles), son pretender, por una parte, fomentar la inversión privada interna en sectores productivos y, por la otra, mantener (al mismo tiempo) un fuerte control al aumento general de la inflación (lo cual se logra, entre otras medidas, a través de tasas de interés elevadas). En cuanto a medios que contradicen los fines buscados por la política, un ejemplo clásico es el de las políticas de descentralización que, generalmente, son impulsadas por un fuerte aparato estatal centralizado a través de criterios centralistas.

Finalmente, puede decirse que los enfoques clínicos tienden a explicar todos los efectos de la política en forma deductiva y racional. El riesgo para el evaluador en este caso es que, al asumir los valores intrínsecos de la política pública estudiada (los que se buscan al especificar los objetivos), puede convertirse en un abogado o defensor acrítico de la política. Por ejemplo, una evaluación clínica del incremento al gasto federal en educación en México durante la década de los años sesenta puede concluir que los objetivos verdaderos aunque tácitos fueron relativamente satisfechos: se mejoró la relación entre el gobierno y el sindicato de maestros al reducirse los riesgos de ingobernabilidad de ese gremio, y así apaciguar, en el mediano plazo, las demandas que amenazaban con crear un foco inquietante de desestabilización en el régimen político. Sin embargo, si bien los responsables de la política educativa de entonces cumplieron sus objetivos reales, no avanzaron mucho en cuanto a los objetivos presentados en el discurso que pretendía justificar esa política: la expansión y mejoramiento de los servicios educativos.²¹

El enfoque normativo busca evitar la visión acrítica del enfoque anterior: el evaluador no fundamenta sus criterios en los objetivos propios de la política pública a evaluar sino que “sustituye esos criterios por otros que escoge él mismo y que le servirán de referencia para medir los resultados observables de la política”.²² El hecho de que la evaluación se efectúe con criterios valorativos propios (o exteriores a la política evaluada) y que a partir de ellos se establezca la calificación de la política, puede deberse principalmente a tres razones: 1) que los objetivos del programa público no hayan podido ser definidos de manera clara

²¹ A propósito de las difíciles relaciones entre el sindicato de maestros y el gobierno mexicano, debe verse el artículo de Alberto Arnaut, “La federalización de la educación básica y normal (1978-1994)”, *Política y Gobierno*, México, CIDE, vol. 1, núm. 2, segundo semestre de 1994, pp. 237-274.

²² Mény y Thoenig, *op. cit.*, p. 290.

y precisa, lo cual obliga al evaluador a plantear sus propios criterios; 2) que el evaluador no comparta los criterios de los objetivos de la política pública en cuestión, o que tenga unos criterios adicionales y primordiales, sustentados en consideraciones de tipo presupuestal, por ejemplo, y 3) que exista un marco legal regulador de la actuación gubernamental y, para su evaluación, normas de conducta restrictivas para los funcionarios públicos o limitaciones presupuestales, lo cual define claramente los criterios que deben adoptarse para la evaluación. Igual que en la perspectiva descriptiva, y al contrario de la clínica, los criterios de este enfoque son *exógenos*, pues provienen de consideraciones externas a las que norman y justifican la política pública evaluada.

Toda política pública, al buscar modificar o conservar la situación prevaeciente en un determinado sector de la sociedad, está impulsada por varios marcos que encuadran su acción. Su punto de partida es una visión general del mundo, en la que supone distinguir lo ideal y lo indeseable. Considera al mismo tiempo un conjunto de valores oficiales y otros particulares (institucionales) que orientan su actuación de largo plazo. La condicionan simultáneamente un conjunto de intereses propios y utilitaristas y una visión de la necesidad social urgente que dirigen su acción inmediata. Y están también los intereses individuales o de grupo, que pretende afectar, favorecer o perjudicar.

En cualquier caso, toda política pública lleva consigo una idea determinada de la sociedad (ideología) y una serie de valores que le ofrecen, aunque sea en forma mínima, un marco justificativo a sus objetivos y a los medios para alcanzarlos: una política pública trae intrínsecamente una teoría social del cambio o de la conservación de la sociedad. Si, por ejemplo, los objetivos de la política pública en cuestión son favorables a criterios de eficiencia económica por encima de los valores sociales, un evaluador normativo, preocupado por la inequidad, puede establecer distancia frente a las políticas y plantear valores propios para la calificación de la política. En este caso, la evaluación asume claramente una posición crítica y externa a la política evaluada al seleccionar en forma deliberada unos criterios de medición que cuestionan la acción de la autoridad pública. El ejemplo puede ser invertido: una política social puede ser duramente criticada desde el punto de vista de un evaluador con criterios de eficiencia económica. En ambos casos, son inevitables consideraciones de tipo moral.²³ El

²³ Aunque de cierta manera las consideraciones morales siempre existen en los demás tipos

riesgo del enfoque normativo es, evidentemente, un exceso de subjetividad en la evaluación o una flagrante arbitrariedad moralista con la condena, con base en un criterio rígido, de todo lo que se salga de la norma. Estos excesos demeritan, claro está, la profundidad del análisis de la evaluación.

Se puede subrayar que dentro de este tipo de evaluación normativa, y más allá de la pretendida objetividad de los números, se ubican los controles que hacen énfasis en el desempeño económico de un programa de gobierno. Si los criterios de la evaluación de una política pública se basan en normas estrictas de cuentas o estados financieros por medio de análisis de costo-beneficio, y ello se hace con independencia de los objetivos iniciales de la misma, es claro que son parámetros externos los que brindan los términos para rechazar o aceptar prácticas y resultados que se consideran deseables o condenables en la evaluación.

Sin embargo, lo más frecuente es que, más allá de esquemas rígidos, se utilicen criterios múltiples para llevar a cabo una evaluación, a través de las técnicas más adecuadas para esos criterios. Aunque siempre predominan ciertos parámetros (corresponde al investigador determinar su orden), hay cinco criterios que en la mayoría de los estudios dedicados al tema se aluden como puntos de partida generales. El criterio de *eficiencia*, o sea la capacidad de transformar los recursos iniciales en resultados finales de manera poco onerosa (logro de objetivos al menor costo posible);²⁴ el criterio de *eficacia*,²⁵ es decir la capacidad de hacer concretas o reales las metas programadas (logro de objetivos tangibles); el criterio de *igualdad*, que implica el mismo tratamiento a todos los sectores sociales, aunque éstos se encuentren en situación desigual; el criterio de *equidad*, que se preocupa en particular

de evaluación, pues siempre se parte de una referencia particular para medir o calificar los efectos de una política. En una sociedad donde abundan valores diferentes, datos y tiempos cuyas definiciones no son compartidas por todos sus integrantes, evaluar implica, por lo menos, asumir una posición y tomar partido.

²⁴ Dicho criterio se mide generalmente a través del análisis de costo-beneficio, procedimiento derivado de la economía que calcula el beneficio neto de la intervención del gobierno en relación con los costos de dicha intervención.

²⁵ Suele haber confusión entre estos términos, que son todos derivados de la palabra *efficere*, a su vez antecedida por *facere*. Eficiente es el participio activo de *efficere*, que significa producir un efecto. El diccionario de la Academia de la Lengua Española estipula que la *eficiencia* es la virtud y facultad para lograr un efecto determinado, y la *eficacia* es la virtud, actividad, fuerza o poder para obrar. Queda así que la *eficacia* implica una disposición, mientras la *eficiencia* traduce una relación condicionada al efecto. Por su parte, Brewer y deLeon definen a la *eficiencia* de manera muy sencilla: es la relación entre insumos (recursos) y productos (resultados). Y *effectiveness* es la relación entre resultados observables y resultados programados en un periodo determinado. Véase Brewer y deLeon, *op. cit.*, p. 335.

por las situaciones de desigualdad social (en cuanto al acceso a los servicios sociales, por ejemplo), para lo cual integra consideraciones distributivas;²⁶ y el criterio de *pertinencia* o *adecuación*, que se refiere a la capacidad de responder satisfactoriamente a las necesidades o demandas sociales.²⁷

¿Cuál es el lugar de estos criterios dentro de los enfoques descritos? El criterio de eficiencia agrega una consideración normativa a los objetivos de la política pública, pues busca establecer una estimación relativamente objetiva —económica— de los recursos con los que cuenta el funcionario (humanos, organizacionales, económicos, institucionales, etc.), para calcular su desempeño tras la observación de los resultados finales. En cambio, la *eficacia* parte de los objetivos programados y los asocia directamente con los resultados concretos, por lo que es un criterio común en las evaluaciones clínicas. Con estas definiciones, puede decirse, por ejemplo, que una política pública puede ser ineficiente si se emplearon recursos descomunales para su aplicación, pero eficaz, si se alcanzaron los objetivos programados, aunque haya sido a un alto precio.

Mientras que la *eficacia* y la *eficiencia* son criterios prácticos, la *igualdad* y la *equidad* incluyen consideraciones de tipo moral. En el primer caso, se considera que la intervención del Estado no debe privilegiar a grupo de individuos o sector alguno, sino que debe actuar en forma imparcial e igualitaria frente a la sociedad. Por su parte, el criterio de *equidad* no sólo sirve para evaluar la actuación del Estado sino que se considera como un factor que incita y justifica plenamente la intervención del mismo, cuando ocurren desigualdades en condiciones sociales o económicas que favorecen la concentración de la riqueza o del poder en un pequeño grupo social. Por lo tanto, los criterios de

²⁶ La *equidad* es la propensión a dejarse guiar, o a fallar, por el sentimiento del deber y de la conciencia, más que por las prescripciones rigurosas de la justicia: es la justicia natural por oposición a la letra de la ley positiva. Por *igualdad* entendemos aquí que a todos los ciudadanos correspondan los mismos derechos, el mismo trato.

²⁷ Por ejemplo, un programa gubernamental es: 1) eficiente, si genera resultados satisfactorios (políticos, sociales, económicos, materiales, etc.) con escasos recursos (humanos, técnicos, organizacionales, financieros, políticos, etc.); 2) eficaz, si en un periodo de instrumentación prestablecido, alcanza verdaderamente las metas programadas; 3) observa el criterio de igualdad si sus beneficios o servicios se reparten en forma equilibrada entre los grupos sociales afectados; 4) observa el criterio de equidad si genera una mayor justicia social tras la intervención gubernamental, a través de estrategias distributivas, por ejemplo; y 5) pertinente o adecuado, cuando la disponibilidad de los servicios o beneficios del programa es suficiente para la necesidad de los demandantes de la intervención gubernamental. Véanse G. Brewer y P. deLeon, *op. cit.*, 1983, p. 338; y L.F. Aguilar, *La hechura de las políticas públicas*, México, Miguel Ángel Porrúa, 1992, p. 29.

igualdad y de equidad intervienen en el enfoque de la evaluación normativa y, como veremos posteriormente, en el pluralista.

Finalmente, el criterio de pertinencia es en cierto modo el inverso del criterio de eficacia, ya que integra completamente el punto de vista del sector social que demanda la intervención de una autoridad pública para atender alguna cuestión de su interés. Se pregunta si el programa o la acción estatal fue la adecuada para satisfacer las necesidades y demandas sociales. Al tomar en consideración a los grupos afectados, este criterio vuelve a situarse en un plano que es independiente de los objetivos, y al imponer valores de medición exógenos, puede situarse en el enfoque normativo o pluralista.

Ya hemos mencionado que, de alguna manera, todos hacemos evaluaciones. En los países donde la competencia decisiva por el poder tiene como medio esencial las elecciones limpias y democráticas, el evaluador más temido por las autoridades (por lo menos las electas) es el ciudadano que vota. En efecto, una forma elemental—sin mayores pretensiones teóricas— de evaluación normativa es la que ocurre, de una manera general, cuando los ciudadanos califican (o descalifican) las acciones de sus gobiernos. Esta valoración se expresa a través del voto y de las opiniones que recogen las encuestas y los medios de comunicación.

“Cada cabeza es un mundo”, reza el dicho: los individuos y los grupos poseen referencias muy heterogéneas para sus juicios de valor, de acuerdo con sus posiciones sociales, su nivel de educación, sus experiencias personales e ideas políticas, su interés particular y su situación material y física frente a la política. Los criterios dependen de quién los formula, de cuál es su intención y su visión del mundo. Los puntos de partida para una evaluación también dependen de la autoridad pública y de la política de la que se trate. Lo cierto es que los criterios y escalas de valores a partir de los cuales los individuos juzgan la actuación gubernamental son subjetivos, por más que el resultado que se busque medir sea totalmente objetivo. Por ejemplo, no ve igual el resultado de una política cambiaría un exportador que un importador: al primero le favorece un tipo de cambio subvaluado que al segundo perjudica. Toda política pública genera efectos que se entienden como adversos por algún sector social; para toda acción del Estado hay adeptos y opositores.

Por otra parte, ni la información concreta ni la realidad suscitan o determinan el carácter de los juicios en la mayoría de los individuos, sino su cultura. Es esta última la que le proporciona al individuo, en

gran medida, el marco moral e intelectual (el marco que regula sobre su conocimiento, sus apreciaciones, sus emociones, su razón, sus sensaciones, su memoria, su percepción) a través del cual puede elaborar un juicio o una evaluación sobre alguna política gubernamental. En esa medida, los hechos exteriores pasan por el filtro de los esquemas de conocimiento que están culturalmente enraizados en los individuos: al mirar los hechos, el individuo tiende a asimilarlos y comprenderlos con los colores de sus juicios preestablecidos. Si algunos elementos de la realidad vienen a cuestionar su marco de valores y las referencias básicas que usa para el entendimiento, el individuo tiende a rechazarlos, a no percibirlos o a torcerlos en el sentido de sus creencias: prefiere seleccionar las partes que menos cimbran ese marco, las que más le convienen y más consolidan sus prejuicios.

De estos vicios de interpretación busca alejarse el evaluador profesional o el analista de políticas públicas. En ese sentido, el enfoque más ambicioso desde el punto de vista *científico* es el que presentaremos a continuación.

El enfoque experimental parte del supuesto de que es posible descubrir relaciones estables de causalidad entre una política pública y sus efectos sociales. Si funciona la política, se puede extender a otras áreas: esto es, el enfoque considera que es posible postular la existencia de conexiones que pueden ser explicadas lógicamente entre las causas—o las variables independientes, que son las medidas específicas del programa gubernamental— y las consecuencias—o variables dependientes, que son los efectos calculables en el sector social o entorno físico donde se desarrolla ese programa—. Por ello, la evaluación experimental, como una ciencia que establece los criterios de su experiencia, propone llevar a cabo ensayos y experimentos de políticas que muestren, en un sector reducido y controlable, en un lugar delimitado y un lapso preestablecido, los efectos reales, para así probar todas las posibles soluciones y alternativas para escoger después la más conveniente (de acuerdo con los objetivos y valores de la autoridad responsable). A una política social distributiva determinada corresponderá una caída previsible del número de individuos que viven por debajo de la línea de la pobreza, pero el fin es generalizar la política que la evaluación experimental exhibió como la más adecuada.

La estrategia evaluativa experimental ha utilizado mucho el análisis comparativo como herramienta para la elección de varias políticas alternativas. Advierte que si el gobierno adopta un conjunto de medi-

das X , sobre un grupo social A , las consecuencias serán del tipo Y ; si adopta un programa alternativo Z sobre un grupo B , los efectos serán W . Después, confronta analíticamente las diferencias entre los conjuntos (X, Y) y (Z, W) desde diversos criterios (económicos, sociales, institucionales, políticos, etc. —vuelve a intervenir el inevitable factor subjetivo—) y decide qué tipo de política le resulta mejor.²⁸ Para llevar a cabo su análisis, la evaluación experimental distingue entre estrategias de investigación y técnicas de tratamiento de la información. Por las primeras entiende la organización metodológica general de la evaluación, en función de las observaciones concretas (información empírica), que estén disponibles o que sea factible recolectar. Las segundas se refieren a los procedimientos o técnicas para decantar los datos empíricos, así como la organización de su tratamiento, su síntesis y su interpretación.

De acuerdo con los objetivos de la evaluación, con los recursos disponibles y las condiciones contextuales, se pueden distinguir dos técnicas de investigación diferentes. Su “pureza experimental” depende de la presencia o no de un grupo “testigo” o de control, equivalente al grupo que es objeto de la política experimentada, para establecer una comparación sistemática.²⁹ En efecto, el carácter experimental de una evaluación depende del grado de similitud entre el grupo afectado por la política pública y el grupo que se usa como referencia comparativa para medir los efectos de la política, esto es, el grupo de control. Si este grupo es por completo equivalente, es decir, si tiene características que son suficientemente similares a las del grupo tratado (desde el punto de vista estadístico), se habla de experimentación real. Cuan-

²⁸ Un ejemplo clásico de este tipo de evaluación se encuentra en M. Deleau, J.P. Nioche, P. Penz y R. Poincard, *Évaluer les politiques publiques: méthodes, déontologie, organisation*, París, La Documentation Française, 1986. p. 75. Éste consistió en la creación de un “impuesto negativo” —una ayuda económica directa— que le fue aplicado a 725 familias con ingresos muy bajos (grupo experimental) y la existencia de un grupo de control de 491 familias que no recibió este tipo de ayuda. La experiencia, llevada a cabo en los estados de Nueva Jersey y Pensilvania, duró tres años (1969-1972), y tuvo un costo de más de ocho millones de dólares. Se buscaba determinar si ese tipo de ayuda directa a los grupos más pobres hacía variar su situación profesional y condición social. Se mostró que los jefes de la familia beneficiada por el impuesto negativo no disminuyeron significativamente su esfuerzo por mejorar su condición social (búsqueda de empleo o de ascenso). Al mismo tiempo, se registró una disminución más significativa en la actividad profesional de las madres de familia. El experimento ha sido muy criticado por los vicios que se introdujeron, en los planos técnicos y morales, durante su desarrollo.

²⁹ La “pureza” alude aquí al sentido más literal del término “experimentar”: probar prácticamente una cosa, ensayarla, como sucede en los experimentos químicos o en las técnicas de evaluación en epidemiología, donde un grupo de enfermos ingiere una vacuna experimental para comparar sus síntomas con un grupo de control al que no se le aplicó. Véase Schwartz, en Deleau, Nioche et al., *op. cit.*, pp. 93-105.

do se comparan los efectos del programa evaluado con un grupo o sector social que no es suficientemente equivalente, se le denomina evaluación *cuasiexperimental* o experimental impura. Pero procedamos a ampliar estas diferencias.

La evaluación casi experimental le permite al investigador tener una idea muy aproximada de los efectos verdaderos de la aplicación de la política, al reducir considerablemente la incertidumbre ligada a la influencia inevitable de factores externos sobre el sector social sometido al programa gubernamental. La existencia de dos grupos, el que está sujeto a la acción y el que sirve como referencia para hacer la comparación, brinda la oportunidad de obtener información adicional sobre las evoluciones que suceden de manera independiente de la instrumentación del programa aplicado. El problema con este tipo de evaluación es, precisamente, la selección del grupo de referencia, pues entre sus diferencias iniciales con el grupo que recibe el tratamiento de la actuación gubernamental pueden estar las causas determinantes de lo que el analista interpreta como efectos de la política pública evaluada.

Por otra parte, mientras más grande y complejo es el grupo seleccionado, más se filtran en la evaluación “impurezas” que se originan en las diferencias de la naturaleza de los dos grupos o sectores comparados. Este tipo de complejidad sucede cuando se evalúa, por ejemplo, una estrategia de desarrollo urbano al comparar dos ciudades, o bien al comparar dos naciones, donde la complejidad y diversidad de los actores y contextos nubla la nitidez de la interpretación por comparación. Si el punto de referencia es malo, el análisis por fuerza es inadecuado: aquí entramos en problemas ligados a la estrategia del análisis comparativo, según el cual, para llevar a cabo una comparación válida, no basta comparar las propiedades similares entre dos cosas, sino que debe cuidarse que los objetos cuyas características se comparan pertenezcan a la misma clasificación general: deben ser casi equivalentes. Lo difícil resulta, por lo tanto, escoger un grupo de referencia que realmente cumpla con los requisitos mínimos del análisis comparativo. Por otra parte, debe cuidarse que el grupo o sector de referencia no sea objeto de otra política pública que se estuviera aplicando simultáneamente.

Así, la simple observación de grupos que no son equivalentes tiene una pertinencia limitada para deducir relaciones causales, puesto que es muy difícil asegurar el grado de comparación entre los grupos. Para salvar estas limitaciones, cuando las políticas públicas están claramente identificadas y pueden ser cuantificables, se puede recurrir

a estimaciones econométricas para evidenciar el peso respectivo de las diferentes variables explicativas. Pero ahí nos encontramos nuevamente con una serie de dificultades técnicas, lo cual ha hecho que este método sea poco utilizado.

En el mundo de la evaluación de políticas públicas, sólo hay un método que es en verdad *experimental*, y que es, por lo tanto, muy costoso de realizar y extremadamente difícil de lograr: se trata de un experimento social, donde la clave está en el método de selección de los grupos de aplicación de la política y el de referencia o de control. Esta evaluación es ideal para poner a prueba un programa público (verificar sus consecuencias): se instala un impuesto municipal (el de radicación, por ejemplo)³⁰ en algunos municipios (sucedió en el Estado de México desde 1993) y se observa escrupulosamente el comportamiento de las finanzas públicas en municipios similares; si los resultados son satisfactorios (desde el punto de vista de los objetivos buscados, que son, en este caso, fortalecer la hacienda municipal con recursos propios), se puede extender la experiencia hacia el resto de la entidad. Pero los grupos de referencia y de control deben ser muy parecidos desde el punto de vista estadístico (presentar las mismas características) y, condición esencial, deben ser escogidos en forma estrictamente aleatoria. A partir de una serie de grupos que son equivalentes, se escoge por sorteo los grupos que servirán para la comparación, el grupo de prueba y el de referencia.³¹ También debe prestarse mucha atención a los criterios de control de las variables, que deben ser definidos en forma clara sin dejar lugar a confusiones; a la forma de medición, que debe ser válida, confiable y precisa, y al control de los factores externos sobre el área de observación (por lo que los términos temporales y espaciales deben quedar perfectamente definidos). Se trata de hacer criterios de evaluación lo más objetivos posible. Por ello son muy pocos los programas que pueden ser objeto de la verdadera evaluación experimental, en especial los que intervienen en sectores sociales o lugares geográficos extensos, o programas muy complejos que se desenvuelven en contextos variables.

³⁰ Impuesto que grava actividades económicas (financieras, comerciales, servicios, etc.) de acuerdo con la dimensión del local utilizado.

³¹ Existe una literatura abundante sobre la manera de constituir los grupos o "muestras sociales" equivalentes. En la práctica, se ha desarrollado a partir de trabajos experimentales en agricultura y evaluaciones de tratamientos médicos. Un ejemplo lo proporciona Daniel Schwartz, "Les techniques d'évaluation en épidémiologie", en Deleau, Nioche, Penz *et al.*, *op. cit.*, pp. 93-105.

Gary Brewer y Peter deLeon sostienen que, en realidad, la calidad de la evaluación dependerá tanto de las habilidades profesionales del investigador como de su juicio, que comienza a emplearse desde que se escogen las herramientas y técnicas para llevarla a cabo.³² Ya se ha dicho, los objetivos expresos o implícitos de la evaluación tendrán un peso determinante en su forma, operación y resultados. Por su parte, Eric Monnier invierte la discusión sobre la evaluación de la siguiente manera: la utilidad de los diferentes enfoques no se encuentra en sus criterios metodológicos sino en la credibilidad de la evaluación ante los actores sociales que están involucrados en el programa público en cuestión. Esa credibilidad, explica este autor, no depende tanto del rigor técnico de los instrumentos utilizados sino de la "legitimidad" que los objetivos de la evaluación y el procedimiento tengan frente a los destinatarios de ella. Al partir de la idea de que la evaluación no es ni puede ser un ejercicio neutro y exterior a las relaciones de poder involucradas en el programa gubernamental observado, sino que es un suceso político que forma parte integral del contexto político del programa, Monnier propone un enfoque *pluralista*, que hace participar a los actores sociales en la conducción de la evaluación a fin de reforzar su credibilidad.³³

El enfoque pluralista parte de una concepción del mundo diferente de la que sustenta el enfoque experimental. Para este último, todo lo verificable con experiencias es verdadero (posee una verdad intrínseca): el investigador o evaluador parte de una teoría general y predice consecuencias factibles (sus hipótesis). Si las observaciones de la experimentación las confirman, entonces la teoría es considerada como válida: la verdad se considera objetiva, factible y constituye una propiedad intrínseca (no convencional). Para el enfoque pluralista, la realidad no puede reducirse ni aislarse, y la evaluación debe hacerse en el contexto de lo que se observa: cualquier pretensión de objetividad es tan inútil como falsa. El analista debe situarse dentro del sistema que evalúa y asumirse como un participante más, pues el simple re-

³² G. Brewer y P. deLeon, *op. cit.*, p. 354.

³³ Eric Monnier desarrolla la idea de la evaluación *pluralista* en *Évaluations de l'action des pouvoirs publics*, París, Economica, 1987. Este autor a su vez habla sobre las siguientes lecturas: D. Greenberg y P. Robins, "The Changing Role of Social Experiments in Policy Analysis", *Evaluation Studies Annual Review*, vol. 10, 1985. Por otra parte, M. Scriven, "The Methodology of Evaluation", en Tyler, Cagne, Scriven (eds.), *Perspective of Curriculum Evaluation*, Chicago, Rand McRally (Monograph Series on Curriculum Evaluation, núm. 1), 1967.

gistro y observación de las acciones, hechos y efectos de la política pública evaluada, suscita su variación.³⁴

En el enfoque pluralista, la verdad tiene sentido en relación con el sistema de referencias del observador y con el entorno que observa: no hay, por lo tanto, una sola verdad, sino varias verdades que coexisten. Si en el enfoque experimental el énfasis se pone en la importancia de los cambios verificables concretamente, en la objetividad y en el rigor de los criterios para evaluar, en el enfoque pluralista la calidad científica depende de la estructura racional que organiza las informaciones producto de la observación. Al considerar que el entorno cambia a través del tiempo, el criterio fundamental de la evaluación pluralista es la pertinencia o la validez de las observaciones dentro de su contexto, en lugar de su factibilidad intrínseca.

Por ello, la evaluación pluralista, al buscar interpretar los efectos específicos de la política pública, requiere de la reconstrucción (o conocimiento) de los diferentes objetivos de todos los actores sociales involucrados: autoridades responsables, elementos del aparato ejecutante, grupos sociales demandantes de la intervención gubernamental y grupos sociales afectados por la acción pública. La evaluación pluralista adopta los puntos de vista de varias categorías de sujetos, grupos o personas interesados por el programa de la autoridad. Considera las diferentes teorías de acción y busca comprender los problemas prácticos que cada uno se plantea a partir de la comprensión de sus actos.

Toda política pública, al ser aplicada, crea un sistema de acción intercomunicado que involucra a las autoridades responsables, a los sectores sociales afectados deliberada o accidentalmente (efectos de rebote), a los cuerpos burocráticos y administrativos que instrumentan la acción, a los sectores de opinión que se interesan en el proceso de la actuación pública, a los grupos que ven en esa actuación una oportunidad de mejorar y a los que la miran como amenaza para sus intereses,

³⁴ El enfoque experimental se origina en la tradición del *individualismo* u *hombre sociológico* de R. Dahrendorf, que comprende a la sociedad como un conjunto de individuos que pueden, en última instancia, considerarse aislados unos de otros (aunque el hombre sociológico no es lo mismo que el hombre económico: el primero puede ser guiado por la fuerza de la costumbre por encima de sus preferencias; no tiene, en muchos casos, una opción clara para su racionalismo, que no es perfecto). Esta visión del mundo se opone a la tradición del *holismo* griego (del griego *holos*, todo), que entiende al universo como un conjunto indivisible por ser un sistema en constante evolución y constituido por elementos esencialmente interdependientes en su interior. Se puede ver a Michel Lallement, *Histoire des idées sociologiques*, París, Nathan, 1993, t. II, p. 192.

etc.³⁵ El evaluador pluralista, a través de sus procedimientos y de sus instrumentos técnicos de evaluación, se integra a la naturaleza del sistema social estudiado para descubrir la influencia específica de la política pública.

Monnier ilustra lo anterior cuando sugiere que su papel es equivalente al de un psicoanalista de la sociedad: el analista pluralista informa constantemente sobre sus percepciones a los actores sociales para que éstos mejoren sus conocimientos, reconsideren sus decisiones y puedan simular posibles consecuencias de sus acciones. El evaluador interpreta y revela las verdaderas tramas sociales (los intereses que están en juego); al mismo tiempo, compromete a los protagonistas en un procedimiento colectivo para entender mejor los sucesos sociales. El evaluador es a la vez un metodólogo y un mediador cuyo papel es integrar y comprender la naturaleza del conjunto de las relaciones y de los intereses en juego para ayudar a especificarlos y clarificarlos. Al cumplir con ese papel, el evaluador pluralista se convierte en un conciliador, “un actor social activo que armoniza a los diferentes grupos y partidos involucrados, y descubre puntos de acuerdos y consensos entre los participantes”.³⁶

Su procedimiento racional es inductivo y permite la definición de un organigrama con varias etapas sucesivas en las que se adapta permanentemente la evaluación, por el hecho de cuestionarse una y otra vez sobre sus propios criterios, sus técnicas, sus objetivos y sus resultados. Veremos ahora la manera de proceder en la evaluación pluralista y los seis (por lo menos) pasos a seguir para llevarla a cabo.

³⁵ Esta concepción de las políticas públicas que estructuran un sistema de acción proviene de las concepciones teóricas sobre las “redes o conexiones de política” (*policy network*). Este concepto fue elaborado, entre otros autores, por Katzenstein, quien lo utilizó para estudiar las influencias del sector público en la formulación de la política exterior de Estados Unidos. A su vez, H. Hecho y A. Wildawsky, en su estudio sobre el Tesoro británico (*The Private Government of Public Money; Community and Policy inside British Politics*, Londres, MacMillan, 1974), introdujeron el concepto de *policy community*, a través del que propusieron que la aplicación de una política pública no podía ser considerada como una fase aislada del resto del proceso (pero pusieron el acento en la fase de la instrumentación de la política): funcionarios y organizaciones privadas, interesados por los mismos problemas, comparten opiniones y actúan en consecuencia. Por lo tanto, el analista debe reconstruir la *red de la política pública* o la *comunidad política* para identificar a todos los actores que participan efectivamente en la formulación y activación de la política pública, los papeles y posiciones de cada cual, las comunicaciones, los canales de influencia, los orígenes de su manipulación, etcétera.

³⁶ Véase Monnier, *op. cit.*, p. 113. Y un resumen de la evaluación pluralista por este autor se encuentra en “Objetivos y destinatarios de las evaluaciones”, *Documentación Administrativa*, *op. cit.*, pp. 131-157.

1) En primer lugar, se forma una comisión de trabajo que integra a todos los participantes interesados en la política pública en cuestión.

2) La comisión tiene como tarea encuadrar lo que está en juego en la evaluación (¿a los intereses de quién sirve?); define después su utilidad de acuerdo con esas prioridades y precisa sus objetivos (es decir, descubre su posible utilidad a partir de las preguntas iniciales y determina si vale la pena hacerla).

3) Se define después con claridad el carácter de la evaluación: aquí, se identifican el sistema de referencias y los criterios de evaluación, y se establecen los valores involucrados en la medición.³⁷

4) Se escoge enseguida la metodología, de acuerdo con una guía de diferentes modelos (Monnier dice paradigmas). Ya se ha dicho, la elección de la técnica responde a diversos criterios dependiendo de la situación de los evaluadores. Se eligen métodos de recolección de información y formas de tratamiento y análisis. Debe reconocerse que cada modelo indica al evaluador lo que es “importante”, legítimo y razonable, y le señala prioridades; inevitablemente, cada paradigma incluye un riesgo de prejuicio.³⁸

5) Se pasa enseguida a la instrumentación de la evaluación. Es una serie de ejercicios en los que se tienen en cuenta los criterios de adaptación de los métodos a las preguntas de partida, la credibilidad de los resultados y la factibilidad de la evaluación dentro de las restricciones de presupuesto. Monnier propone una evaluación que sea adap-

³⁷ Al seguir a M. Scriven (“Merit versus value”, *Evaluations News*, 1978), Monnier identifica un cuadro general de cinco sistemas de valores diferentes en cualquier marco institucional: 1) los valores declarados, reivindicados por la retórica de la institución; 2) los valores de hecho, que le son propios a la institución (producto de los usos y costumbres de su funcionamiento, por ejemplo); 3) los valores inmediatos de la institución —valores intrínsecos—, desde un criterio estrictamente utilitarista, y cuando la acción está guiada por la intención de proteger o favorecer intereses; 4) los valores contextuales, que están definidos a partir de la situación y la posición en la que se lleva a cabo la evaluación (intereses políticos en juego, restricciones presupuestales, marco legal limitativo, etc.), y 5) los valores normativos, que rigen al comportamiento que se considera como ideal.

³⁸ Los seis paradigmas de Monnier son: 1) *el modelo lógico*: partiendo de axiomas, todo enunciado es verdadero si se puede hacer su demostración lógica; 2) *el modelo experimentalista*: es verdad todo lo que sea verificable en forma experimental; 3) *el modelo endosistémico* (o pluralista): el observador reduce el fenómeno observado (pues lo real es un todo, un sistema irreductible) y la verdad sólo tiene sentido en relación con el observador y su entorno; 4) *el modelo del experto*: la verdad es inducida por la experiencia humana de la realidad construida desde el pensamiento respetado y reconocido de un experto; 5) *el modelo de pretoria*: donde la verdad se deduce de la confrontación de diferentes puntos de vista, del debate sobre tesis opuestas y de la confrontación de pruebas contradictorias. Son prejuicios a través de los cuales se puede llegar a un juicio que se asume como verdadero o definitivo (son las evaluaciones que llevan a cabo comisiones parlamentarias, fruto del debate entre fracciones políticas); y 6) *el modelo demográfico*: la verdad es concebida a través de un examen macroscópico de los fenómenos socioeconómicos.

table y creíble, y que evolucione a lo largo de su instrumentación. El trabajo del evaluador significa una sucesión de análisis efectuados a partir de puntos de vista numerosos y complejos, para decantar, separar y clasificar lo más posible los diferentes aspectos de la evaluación.

6) Finalmente, la comisión pasa al análisis de los resultados (todos los actores participan) y decide sobre la utilización que se les dará (corrección o adaptación de la política, etc.), y sobre los medios a través de los cuales les dará difusión (a nivel institucional y público).

Una particularidad de este enfoque es la importancia que se otorga a la definición de los criterios de evaluación. El evaluador debe preguntarse si los resultados obtenidos de la evaluación pueden ser difundidos, por lo menos dentro de la institución que haya llevado a cabo la evaluación. Si la respuesta es negativa, debe volverse a empezar el proceso de definición de criterios, hasta que éstos sean compatibles con las condiciones de difusión interna (y de preferencia también externa) sin que se incurra en graves conflictos de interés. Por ello, recuerda Monnier, el papel del evaluador es también el de crear consensos a partir de la participación general de quienes tomaron parte activa en el programa a evaluar.

La dificultad de llevar a cabo una verdadera evaluación pluralista es la misma, toda proporción guardada, que la de construir instituciones democráticas puras y verdaderamente pluralistas: suena muy bien en teoría. Pero los retos de practicarla con apego a cada uno de sus supuestos básicos son enormes, pues en el mundo real dichos supuestos encuentran obstáculos muy difíciles de librar. Al ironizar sobre lo impracticable que resultaría una democracia pura en un entorno pluralista, muchos autores aluden simplemente a la absoluta parálisis gubernamental que se tendría si todas y cada una de las decisiones públicas tuvieran que ser sometidas a referéndum después del debate de rigor: jamás nada podría ser decidido, pues todo lo público suscita controversia y la unanimidad es utópica en un contexto de pluralidad. En forma análoga, uno se puede preguntar: ¿cómo podría formarse y funcionar una comisión que integrara a todos los aludidos por la política pública en cuestión? ¿Cómo hacerlo cuando se trate de una política muy compleja y de envergadura nacional, como por ejemplo, las reformas al poder judicial emprendidas en 1995 por el gobierno del presidente Ernesto Zedillo? ¿Cómo suponer que los puntos de vista de todos los actores, los más y los menos influyentes, los más y los menos afectados, los ejecutantes y los beneficiados, van a tener un peso equiva-

lente en la hechura de la evaluación? ¿Cómo creer que no va a predominar el punto de vista de quien finalmente cubre los gastos de la evaluación?

Todo lo anterior nos lleva inevitablemente a las siguientes dos conclusiones: la primera, que no hay un enfoque de evaluación que resulte mejor que otro. La evaluación de una política pública depende de su carácter y del sector al que se dirige; de la naturaleza de las autoridades públicas encargadas de la política estudiada; de los recursos (humanos, materiales, técnicos, tecnológicos, etc.) y del tiempo con el que se cuenta; de la independencia del evaluador respecto al patrocinador de la evaluación, etc. ¡Es una lista muy larga de factores diversos! Para algunas políticas, los investigadores han tenido que conformarse con un procedimiento descriptivo, mientras que en otras, han podido llevar a cabo un costoso programa de experimentación de política pública (en realidad, los ejemplos en este sentido están sobre todo en los países desarrollados).

La segunda conclusión es que la utilidad de una evaluación no debería depender solamente de un patrocinador dispuesto a verificar y estudiar los efectos de una política pública. Sus resultados deberían ser públicos para que los afectados en la sociedad pudieran controlar mejor los actos de las autoridades y, ocasionalmente, si las fallas lo ameritaran, poder llamar a esos responsables para que rindieran cuentas de sus errores, de sus excesos o de sus abusos. Para que esto sea posible, es necesario que la evaluación sea patrocinada por un órgano público diferente al que ejecuta la política, o que el marco legal e institucional obligue en forma efectiva al ejecutor a que ofrezca cuentas y explicaciones de los resultados de su actuación. De estos problemas, que se resumen en la necesidad de contar con instituciones que propicien dentro del sistema político la evaluación y el control real sobre el poder ejecutivo, nos ocuparemos en la segunda parte de este trabajo.

Terminamos esta primera parte con un resumen esquemático que sintetiza los rasgos principales de cada uno de los enfoques descritos; su pretensión es identificar someramente los principales rasgos distintivos de cada enfoque, para también plantear sus diferencias fundamentales.

Para cada enfoque se enlistaron los objetivos esenciales de la evaluación, y la relación de éstos con los objetivos generales de la política pública evaluada. En muchos casos, ya se vio, los objetivos de la evaluación consisten en verificar si las metas de la política se cumplieron

y medir (cuantificar y calificar) sus efectos sociales y los que ocurren sobre el ambiente en el que opera la política. Son los casos de los enfoques clínico, experimental y pluralista, que no pierden de vista, a lo largo de la evaluación, lo que la autoridad pública buscaba lograr desde el inicio de la instrumentación de su programa. En esos casos, los objetivos de la evaluación y de la política pública están directamente relacionados y, por lo general, la evaluación está subordinada a las metas de la política pública.

En los casos de las evaluaciones descriptivas y normativas, los objetivos de la evaluación son relativamente independientes de las metas de la política pública: el descriptivo saca una fotografía y describe lo que mira, observando con cuidado los cambios concretos tras la aplicación de la política. El normativo busca medir, con su propia escala de valores, los efectos de la política en los rubros que interesan al evaluador (si es la rentabilidad, por ejemplo, el evaluador puede concluir felizmente que una política pública es muy rentable, aun cuando sus efectos sociales sean desastrosos). El evaluador normativo se pone sus lentes propios (legales, morales, económicos, etc.) y a partir de sus criterios particulares, califica y cuantifica los efectos de la política.

Es posible que coincidan, en forma expresa, los criterios y los objetivos de la política pública y los de la evaluación normativa. Esto es frecuente: cuando se quiere dar a una acción pública un sustento legal y se le evalúa con criterios jurídicos, o cuando se plantea una reducción presupuestal y se evalúa con criterios de eficacia y efectividad, o cuando se evalúan con criterios comerciales los efectos de una reducción de tarifas arancelarias en el sector externo. Entonces, criterios, valores y marco normativo de la política son secundados por los criterios, valores y normas de su evaluación. Existe una lógica en ello: algunas políticas públicas tienen objetivos y criterios que sólo pueden ser medidos con la misma vara. Sería incongruente que no se midieran los efectos del Tratado de Libre Comercio de América del Norte (TLCAN) en la actividad económica de México, y para ello son necesarios criterios de evaluación económicos. Pero sería insuficiente si no se consideraran sus consecuencias sociales y políticas que, por otro lado, no son las mismas en el estado de Sonora que en Oaxaca, en los medios de comunicación que en el sector educativo, en los patrones de consumo que en los patrones culturales. Es deseable que las políticas públicas sean sometidas también a criterios y escalas de evaluación alternativos, pues tienen, de manera invariable, efectos que se dispersan mucho más allá de los que declaran oficialmente sus defensores.

Al aludir a los criterios, en el resumen esquemático hemos querido distinguir el procedimiento general del enfoque, su modo de razonar y su punto de referencia. Identificamos, de una forma general, dos métodos de razonamiento: aquél que saca consecuencias lógicas a partir de una idea, un principio o un conjunto de hipótesis, esto es, el método deductivo; y el que infiere una cosa a partir de los hechos particulares o de las leyes que los rigen, el método inductivo. En términos simplificados, la deducción lleva de lo abstracto (del mundo de las ideas, la teoría) a lo concreto (los hechos, la confrontación con la realidad); y la inducción hace el trayecto contrario, de la observación a la teoría.³⁹ Un razonamiento inductivo tiene como punto de partida la observación de los efectos concretos de la política pública, y como punto de llegada la interpretación de su significado en el sector social afectado; por su lado, el razonamiento deductivo parte de algunos principios o enunciados para mirar, bajo sus criterios, los efectos de la política, y así establecer sus conclusiones. La evaluación descriptiva parte de los hechos y sigue después el razonamiento inductivo, al igual que los enfoques experimental y pluralista. Las evaluaciones normativas o clínicas usan el método de razonamiento deductivo. Pero cada uno de los cinco enfoques tiene puntos de referencia (para fundamentar sus criterios) distintos; por eso, si la misma política pública fuera evaluada por cada uno de los enfoques que hemos venido repasando, las conclusiones de cada uno serían muy probablemente diferentes.

Ello está también ligado al hecho de que lo que cada enfoque entiende como verdadero (bueno, mejor, etc.) o considera como deseable (o condenable), es decir, los términos que cada uno tiene para valorar la política pública (en sus metas, procedimientos, prácticas y efectos), varían de acuerdo con los objetivos y los criterios de evaluación. En el enfoque clínico, en última instancia, corresponde a las autoridades la definición de los valores para medir las consecuencias de la política pública; en el enfoque normativo toca a los evaluadores decidir sobre las normas y principios para cuantificar y calificar a la política. Los valores son externos en el caso del enfoque descriptivo; dependen de todos los participantes en el del pluralista, y son proporcionados por la experiencia en el enfoque experimental.

³⁹ Pascal Engel define a la inducción como una inferencia no deductiva, que pasa de premisas particulares a una conclusión general que se asume como probable. Aquí, queremos solamente subrayar que el razonamiento se invierte entre el método deductivo y el inductivo, lo cual tiene consecuencias en la definición de los criterios de evaluación. Véase P. Engel, *La norme du vrai. Philosophie de la logique*, París, Gallimard, 1989, p. 476.

Resumen esquemático de los rasgos principales de los enfoques de evaluación

1. Evaluación descriptiva

Objetivo:	Es independiente de los objetivos de la política pública. Registra cambios concretos después de la aplicación del programa gubernamental. No explica las causas de los efectos ni mide los niveles de variación.
Criterio:	Razonamiento inductivo. Compara la realidad antes y después de la aplicación de la política. El punto de referencia es la situación anterior a la instrumentación de la política.
Técnicas:	Muy variadas: elaboración de muestreos, investigación por encuestas y circulación de cuestionarios, análisis estadísticos, observación participante o directa, análisis de contenido (informes gubernamentales), construcción de escalas o de modelos matemáticos, etc. Y el estudio de caso, cuando éste se limita a la descripción.
Valores:	Es verdadero todo lo que se puede registrar como cambio. Hace un examen general del lugar o sector en donde se instrumentó la política pública e identifica los cambios concretos, sin calificarlos (no caracteriza lo deseable, lo superior, etcétera).

2. Evaluación clínica

Objetivo:	Mide la efectividad de la política pública con base en los objetivos de la misma y la califica con sus normas particulares. Asume los objetivos de la política pública.
Criterio:	Razonamiento deductivo. Confronta los objetivos de la política pública con los resultados concretos e interpreta las fallas. Mide la efectividad con base en la consecución de objetivos, que son su punto de referencia.
Técnicas:	Sobresalen tres: la encuesta estadística; la <i>evaluación sumaria</i> (antes y después); y la <i>evaluación formativa</i> (seguimiento progresivo).
Valores:	Los propios de la política evaluada: lo verdadero o deseable es definido por la autoridad responsable.

3. Evaluación normativa

- Objetivo:** Califica la política pública con base en normas externas. El evaluador puede definir sus propios objetivos, con relativa independencia de los objetivos oficiales de la política pública (aunque es frecuente que coincidan).
- Criterio:** Mide la eficiencia, eficacia, equidad y pertinencia con base en las definiciones y reglas propias del evaluador. El razonamiento es deductivo: lo deseable/condenable es consecuencia del pensamiento del analista. El punto de referencia es la escala de reglas propuestas por él mismo.
- Técnicas:** Todas las técnicas están sujetas al criterio del evaluador. A través del análisis de la técnica en particular, se puede descubrir el criterio de evaluación.
- Valores:** Son propios del evaluador, con normas y principios establecidos formalmente y con anterioridad a la evaluación. A partir de esos principios, lo verdadero o deseable se puede deducir en forma lógica.

4. Evaluación experimental

- Objetivo:** Controla *científicamente* los efectos concretos de una política pública a partir de sus objetivos oficiales.
- Criterio:** Comparación (estadística, descriptiva, analítica, etc.) del grupo afectado por la política pública con un grupo de referencia o de control. El razonamiento es inductivo: se observan los efectos de la política a través del análisis del comportamiento de los grupos de control y tratado. Su punto de referencia comparativa es el grupo de control.
- Técnicas:** Dos tipos principalmente: comparaciones con un grupo de control equivalente, escogido en forma aleatoria, y con un grupo de control no equivalente (desde el punto de vista estadístico).
- Valores:** Es verdadero todo lo verificable a través de la experiencia: si las observaciones confirman las expectativas teóricas de la política pública, se le considera como válida (sin hacer consideraciones morales).

5. Evaluación pluralista

- Objetivo:** Establece un análisis colectivo y permanente para adaptar y mejorar la aplicación de la política pública con el acuerdo de todos los involucrados y a partir de los objetivos generales de la política.
- Criterio:** Razonamiento inductivo: a partir de los resultados que se van observando, todos los participantes contribuyen —noción de consenso— para definir los criterios de efectividad, eficiencia, equidad y adecuación. El punto de referencia varía, es flexible y resultado del consenso.
- Técnicas:** Una comisión pluralista escoge la técnica, de acuerdo con las necesidades del caso y con base en seis modelos: el lógico, el experimentalista, el endosistémico (o pluralista), el modelo del experto, el modelo de pretoria y el demográfico.
- Valores:** Definidos por los actores, a partir del contexto: lo verdadero y deseable sólo tiene sentido en función del observador y su posición frente a los demás y a su entorno.

Segunda parte. La evaluación de las políticas gubernamentales en nueve países

Presentación

Desde el punto de vista de la disciplina del análisis de la política pública, la investigación sobre los métodos de evaluación como una de las fases fundamentales de la vida del programa gubernamental es perfectamente necesaria y justificable por el solo hecho de contribuir a su perfeccionamiento y a crear conocimiento.

Por medio de la evaluación, por ejemplo, se busca comprender por qué una política de ayuda directa a los grupos de bajos ingresos viene acompañada de efectos inesperados en los planos social, económico y político; por qué decisiones importantes que se toman en la cima del gobierno se pierden sin remedio por las inercias de los cuerpos burocráticos que tenían el encargo de ponerlas en práctica; por qué la información que alimenta una decisión llega al responsable seriamente manipulada; por qué actuaciones gubernamentales que aparecen como totalmente irracionales a los ojos del público tienen su razón de ser cuando se miran desde la perspectiva de otros intereses no explícitos. La evaluación puede ayudar a profundizar en el conocimiento de la

manera como está funcionando el Estado, a comprender mejor sus lógicas de operación, actuación y evolución, a descifrar más atinadamente la naturaleza de las organizaciones burocráticas, y a interpretar mejor la compleja relación entre los sectores públicos y los diversos sectores de la sociedad.

Pero la evaluación no brinda sólo esas oportunidades para ensanchar el conocimiento. Para muchos, el dominio del conocimiento también puede tener otro fin: establecer un instrumento de control. En este sentido, el motor del desarrollo de las evaluaciones se encuentra en los esfuerzos de crear controles de las políticas gubernamentales y asegurar *accountability* —responsabilidad— de los funcionarios encargados. La necesidad de evaluar está también en la necesidad de mejorar el uso de los recursos escasos; en la necesidad de mejorar el desempeño público a través de mayor flexibilidad y efectividad; y, desde luego, en el interés de los gobernantes de mantener y acrecentar su poder sobre las áreas subordinadas de la administración. El conocimiento a través de la evaluación es, en esta segunda acepción, una rica fuente para el fortalecimiento del poder político, pues se convierte en algo práctico, con carácter instrumental: una autoridad evalúa con el fin de inspeccionar a los ejecutores de un programa de gobierno.

Una evaluación de algún programa educativo ordenada por la Coordinación del Gabinete de la Presidencia de la República puede poner en entredicho al secretario de Educación. O bien, este último, al evaluar los avances en materia de descentralización, puede poner en aprietos al delegado de la SEP de algún estado. En fin, es mediante la evaluación como algunas entidades del Estado (en cada nivel de gobierno y entre las mismas entidades) pueden ejercer efectivamente la supremacía y el dominio sobre otras.

En la mayoría de los casos, son los mismos organismos públicos los que ordenan, patrocinan o llevan a cabo las evaluaciones. Pero en el ideal de los casos, este tipo de controles sobre los ejecutores de las políticas públicas se encuentra en otros sectores del Estado o de la sociedad. Esto ocurre cuando un poder distinto al ejecutivo, generalmente el poder legislativo, tiene a su alcance los recursos materiales, políticos e institucionales para evaluar programas públicos y controlarle al ejecutivo el uso de los recursos (presupuestales, humanos, organizacionales, etc.). Ello es condición necesaria para que sociedades complejas, con enormes aparatos administrativos, disfruten de gobiernos que respondan de hecho por sus actuaciones, ofrezcan cuentas claras a los ciudadanos y encuentren equilibrios reales al

ejercicio de su poder. Es decir, que en realidad vivan el equilibrio de poderes a través de la *accountability* y los *checks and balances*, en términos de lo que Carlos Fuentes añora para el sistema político mexicano. En las páginas que siguen, estudiaremos las formas en que las instituciones políticas de nueve democracias occidentales han ido resolviendo el problema del control y la evaluación objetiva de los programas de gobierno.

Consideraciones previas a la comparación

Es relativamente reciente que las acciones gubernamentales sean evaluadas de acuerdo con objetivos claramente definidos y utilizando criterios, valores y métodos sistemáticos. Y es más reciente aún, en la década de los años sesenta, que esas evaluaciones formen parte de los programas oficiales o se integren a la vida institucional del Estado en algunos países. Sin embargo, las políticas de evaluación han alcanzado distintos grados de desarrollo según el caso.

Un estudio comparativo de la evaluación en ocho países, dirigido por Ray C. Rist,⁴⁰ de la Oficina General de Cuentas del Congreso de Estados Unidos (General Accounting Office), muestra que las diferencias atañen no sólo el nivel de análisis, los instrumentos técnicos, la profesionalización y el presupuesto asignado, sino también a la regularidad con la que se llevan a cabo las evaluaciones, así como el grado de su institucionalización en el aparato del Estado y en la vida política del país. En algunas culturas político-administrativas, la evaluación puede darse como un hecho frecuente; en otras, es una novedad o apenas un proyecto. Hoy en día las evaluaciones son todavía una actividad casi exclusivamente pública, pues en la mayor parte de los casos corresponde efectuarlas a organismos gubernamentales, aunque la creciente influencia de los medios de comunicación y la multiplicación de organismos privados que venden sus servicios especializados han permitido que la evaluación de la actuación pública comience a desarrollarse en sectores civiles más o menos independientes del Estado.

Esta segunda parte se ocupará de describir las formas de evaluación en nueve países desarrollados: Alemania (República Federal Alemana), Canadá, Dinamarca, Francia, Estados Unidos, Noruega, Reino

⁴⁰ Lo citamos con anterioridad, es Ray C. Rist (ed.), *Program Evaluation and the Management of Government*, New Brunswick, Transaction Books, 1990, 184 p.

Unido, Suiza y los Países Bajos.⁴¹ A través de un marco comparativo que se establecerá en principio, y que se interesa en las diferentes formas de organizar una evaluación, describiremos los usos particulares que cada país hace de las evaluaciones y analizaremos las condiciones institucionales que han permitido la aparición de las políticas de evaluación en esos países. Por último, presentaremos nuevamente un resumen esquemático de los trazos principales de la evaluación en los nueve casos.

Hemos dicho que al hablar de evaluación nos referimos principalmente a la evaluación de resultados de políticas, en oposición a la evaluación *ex ante* que busca pronosticar los efectos de las alternativas políticas al alcance del responsable de la acción gubernamental. Sin embargo, ya lo vimos por ejemplo con el enfoque pluralista, no se puede considerar que la evaluación de efectos sea estrictamente retrospectiva o *ex post*. En realidad, la evaluación puede iniciar al mismo tiempo que la puesta en práctica de la política y acompañarla en cada una de sus fases, con el objeto de ir midiendo sus consecuencias, para modular así su aplicación, o bien cancelar o modificar su rumbo. Siendo así, la evaluación puede acompañar a la política pública desde el inicio hasta el fin de su vida. A principios de los años ochenta, en Estados Unidos, este tipo de evaluación fue inclusive implantada a través de la creación de dos comisiones gubernamentales que definieron unos criterios generales de evaluación que apuntaban precisamente en este sentido.⁴²

Desde muy temprano, en el proceso de una política pública se puede pedir cuentas a sus formuladores y ejecutantes, identificar efectos y resultados (aunque sean incipientes) y, también, establecer preguntas sobre la responsabilidad legal, social y política de quien tomó la decisión. Pero no hay una sola forma de atacar este problema: como veremos, cada uno de los nueve países analizados tiene su forma par-

⁴¹ R.C. Rist se ocupa de todos los casos menos del francés. Para Francia tomamos el trabajo ya citado de J. P. Nioche y R. Poinsard, "Public Policy Evaluation in France", *Journal of Policy Analysis and Management*, vol. 5, núm. 1, 1985, pp. 58-72.

⁴² En 1981 se instituyó este tipo de evaluación a través del Comité Conjunto de Criterios Generales para la Evaluación Educativa (*Joint Committee for Standards for Educational Evaluation*) y en 1982 con el Comité de Criterios Generales de la Asociación de Evaluaciones de la Investigación (*Evaluation Research Society Standards Committee*). Dichos criterios enlistaron seis estrategias diferentes para conducir evaluaciones de programas: análisis inicio-fin, análisis de posibilidades de evaluación, evaluación de procesos, evaluación de efecto y eficiencia, monitoreo de programa y de problemas, y síntesis de evaluación o análisis de evaluación (*meta-evaluation*, que consiste en evaluar las políticas de evaluación anteriormente ejecutadas). Véase R. C. Rist, *op. cit.*, p. 5.

ticular de resolver, con mayor o menor éxito, las cuestiones de la evaluación de políticas, la vigilancia de la responsabilidad (*accountability*) y las formas de control institucional de las autoridades gubernamentales (*checks and balances*).

Al analizar comparativamente a varios países democráticos, se puede decir que el origen principal de las diferencias en las políticas de evaluación está en las principales características de cada sistema político, así como en los marcos constitucionales que las encuadran.⁴³

R.C. Rist propone analizar los orígenes de estas diferencias al distinguir cinco áreas ubicadas en el marco institucional; para este autor, son éstas las más significativas para explicar las variaciones internacionales.

- 1) El grado de centralización (o fragmentación) del sistema o la organización que se encarga de practicar la evaluación.
- 2) El grado de independencia de las unidades de evaluación respecto a los ejecutores de la política evaluada y al carácter interno o externo de las evaluaciones.⁴⁴
- 3) El destino final (puede ser el poder ejecutivo o el legislativo, una organización social o el público en general) de los resultados de la evaluación y de los informes que formulan sus encargados.
- 4) El acceso a la información sobre las conclusiones de las evaluaciones por parte de los demás actores sociales.
- 5) El nivel de integración de las evaluaciones con la función que desempeña tradicionalmente, dentro de la organización general de la administración pública federal, el auditor, revisor o contador de las cuentas gubernamentales.

Hemos visto en la primera parte que a través de la definición de los objetivos, criterios, métodos y valores de la evaluación se establecen en gran medida los parámetros de lo que resultará de ésta. Pero si los analistas que evalúan están organizados en un grupo centralizado, con una estructura jerarquizada para la toma de decisiones, un orga-

⁴³ Cada sistema político-administrativo tiene sus estilos, instrumentos e instituciones para organizar las evaluaciones, y encuentra en el camino ventajas y obstáculos estructurales que determinan el papel y los resultados de las mismas. Entre las determinantes estructurales podemos contar el marco constitucional, la cultura político-administrativa, las inercias y colusiones burocráticas, la índole de la relación entre el poder legislativo y el ejecutivo, etcétera.

⁴⁴ La distinción es necesaria porque una evaluación externa no garantiza forzosamente la independencia de los evaluadores respecto a los responsables de la política pública.

nigrama de funciones con reglas precisas y una comunicación poco fluida en su interior, los resultados de la misma tendrán que ser forzosamente distintos de los que se obtienen si la organización tiene un poder de decisión distribuido, los papeles de sus miembros dependen de su competencia profesional, sus funciones específicas o de su personalidad, y si existe una comunicación ágil entre sus agentes (evaluación descentralizada).

El grado de independencia de las unidades de evaluación para definir sus objetivos y criterios e iniciar su trabajo es otra consideración importante. Entre otras cosas, ofrece una estimación sobre la capacidad de los evaluadores ante la información que deben recoger, depurar, manejar e interpretar. Si la autonomía del grupo evaluador para establecer su agenda de trabajo es limitada, lo será consecuentemente su capacidad para establecer criterios que se salgan del marco prescrito por quienes tuvieron la autoridad suficiente para precisar cuáles serían las cuestiones evaluadas. Pero sucede lo contrario si las unidades de evaluación cuentan con un grado considerable de independencia y de habilidad para definir sus objetivos, criterios y valores.

En las evaluaciones donde los analistas son internos y los objetivos son definidos por los solicitadores (comanditarios, dice E. Monnier) de las mismas, los estudios tienden a ocuparse de manera especial de las necesidades de información de los responsables de la política pública, que están generalmente interesados por el corto plazo. La posibilidad de plantear cuestiones que se refieran a resultados, efectos sociales y políticos, así como a problemas de pertinencia, equidad e igualdad, trae siempre el riesgo de crear antagonismos entre los evaluadores y los responsables políticos. Que los analistas de la evaluación estén protegidos de la posible indiferencia, la descalificación, el silenciamiento o de las represalias (laborales, presupuestales, administrativas, etc.) por parte de los ejecutores políticos, depende de la independencia y la autonomía de que gocen los evaluadores dentro del marco legal o institucional.

Asimismo, el destino de la evaluación influye sobre la redacción de sus conclusiones, que varían cuando son referidas al funcionario ejecutor, cuando se dirigen a una instancia de control superior, o cuando se remiten a una comisión del poder legislativo. Por lo general, los grupos de evaluadores que informan al mismo cuerpo público que llevó a cabo la política (poder ejecutivo) tienden a llevar a cabo estudios de tipo clínico o bien normativo, haciendo énfasis en el control de la gestión y la eficiencia, y asumiendo los objetivos de los ejecutores y los criterios

del organismo al que le responden. Al contrario, los evaluadores que provienen de comisiones legislativas se centran en sus necesidades particulares (evaluaciones normativas con criterios definidos por un órgano distinto al ejecutor) para la verificación del cumplimiento de objetivos, el control del presupuesto, el cumplimiento de reglas y procedimientos, etcétera.

Otra distinción interesante, señalada en el cuarto punto, atañe a la determinación de los actores que tienen acceso a las conclusiones de las evaluaciones. Algunos países permiten un acceso irrestricto a los informes de sus evaluaciones a casi todo aquél que lo solicita (desde legisladores hasta reporteros de prensa, así como investigadores y estudiantes interesados en el asunto). Otros países tienden a restringir los informes finales de la evaluación a las comisiones legislativas que están jurídicamente facultadas para su uso, y otros sólo permiten a miembros de la administración el acceso a la información proporcionada por las evaluaciones. En algunos casos, las comisiones parlamentarias pueden estar facultadas para hacer del dominio público alguna parte de los resultados o elaborar una versión adaptada para la prensa. En cuanto a los funcionarios del gobierno, suele suceder que las conclusiones se hagan públicas sólo después de una discusión interna, en que se les depura y edita hasta la satisfacción del responsable del área administrativa en cuestión, a fin de que el informe quede bajo control y no presente ningún tipo de sorpresa para la administración. Es muy frecuente, sin embargo, que los ejecutores responsables de la política pública prefieran mantener una difusión muy restringida de los resultados de las evaluaciones que competen a las áreas de su responsabilidad.

Finalmente, otra forma de comparar las distintas maneras que cada nación tiene para resolver el problema de la evaluación es analizar el grado de integración (o de separación) que la evaluación tiene con la función tradicional de las auditorías y la revisión externa de las cuentas públicas. En los nueve casos que revisaremos, la evaluación está de alguna manera asociada con las oficinas generales de auditoría. Si bien es cierto que las lógicas de la auditoría y de la evaluación no son las mismas (a diferencia de la evaluación, la auditoría no proviene de la investigación científica sobre aspectos sociopolíticos),⁴⁵ tienen

⁴⁵ La auditoría parte de una posición normativa que se pregunta si una serie de normas, reglas, condiciones presupuestales, procedimientos y mandatos particulares han sido cumplidos. Subraya el rigor del cumplimiento con criterios claramente preestablecidos. Por el contrario, la evaluación parte de una serie de suposiciones o interpretaciones apriorísticas (hipótesis) para explicar —o por lo menos describir— lo que pasó.

puntos en común: pueden originarse en un análisis retrospectivo que se dirige hacia acciones presentes o pasadas y hacia procedimientos o sistemas organizacionales.

Por otro lado, hay tres factores que han provocado recientemente que las dos actividades se acerquen. En primer lugar, se ha modificado el interés hacia la acción gubernamental: en la actualidad, los gobernados en general, y los miembros de las administraciones en particular (es decir, aquéllos que recurren a la información de la evaluación), no buscan exclusivamente datos que se refieran a cuestiones financieras o económicas (eficacia y eficiencia), piden además información sobre los efectos sociales y políticos de la puesta en práctica de una política pública. Del lado de los gobernados, la cuestión es saber si los gobernantes utilizan con honestidad, limpieza y pertinencia los recursos públicos que son el producto de sus impuestos, por una parte, y si su acción atiende a las demandas de los grupos sociales afectados, con el objetivo último de mejorar su bienestar. Del lado de los propios funcionarios, no resulta suficiente preguntarse si los fondos públicos son gastados de manera correcta, sobre todo en tiempos de escasez presupuestal y de abundancia de obligaciones políticas; es importante saber también si su manejo y su gasto representa una diferencia notable y deseable que justifique su acción y, luego, su repetición.

En segundo lugar, en tiempos recientes se ha desarrollado una convergencia en los métodos que ambas funciones utilizan. La evaluación, lo vimos con anterioridad, puede conceptualizarse casi en su totalidad por los criterios del evaluador y desarrollar métodos estrictamente normativos, al grado que, siguiendo la clasificación de la Primera parte y de acuerdo con la definición del enfoque normativo de la evaluación, podríamos decir que la auditoría no es sino una forma particular de evaluación.⁴⁶ En este sentido, la auditoría ha ido integrando otros criterios que son básicos de la función evaluadora al asu-

⁴⁶ Por ejemplo, en México corresponde a la Cámara de Diputados nombrar a los jefes y funcionarios de la Contaduría Mayor de Hacienda. Esta última examina la Ley de Ingresos y el Proyecto de Presupuesto de la Federación (presentadas cada mes de noviembre) y revisa la Cuenta Pública (presentada en junio de cada año), a fin de “conocer los resultados de la gestión financiera, comprobar si se ha ajustado a los criterios señalados por el Presupuesto y el cumplimiento de los objetivos contenidos en los programas”. El artículo 74 de la Constitución Política, en su fracción IV, agrega que “si del examen que realice la Contaduría Mayor de Hacienda aparecieran discrepancias entre las cantidades gastadas y las partidas respectivas del Presupuesto o no existiera exactitud o justificación en los gastos hechos, se determinarán las responsabilidades de acuerdo con la ley”. Pero la ley no demanda de la Contaduría que lleve a cabo evaluaciones que trasciendan el ámbito fiscal o presupuestario.

mir cada vez más ejercicios descriptivos y, en ocasiones, intentos de explicaciones de causalidad.

El tercer factor que crea lugares comunes entre las dos funciones es de índole práctica: en casi todos los países que veremos, la evaluación se ha desarrollado a partir de las oficinas públicas que llevan a cabo las auditorías tradicionales. En muchos sentidos, y sobre todo cuando se pretende asociar la idea de la evaluación a una forma de control de las autoridades que usan los recursos públicos,⁴⁷ la auditoría se presenta como un campo de estudio ineludible, por las implicaciones políticas que introduce irremediamente (criterios de control normativo, verificación y responsabilidad —*accountability*—) y por las instituciones que en la práctica se ven comprometidas en su instrumentación (en casi todos los países, el poder legislativo —criterio de *checks and balances*).

En los países que veremos, con sistemas políticos democráticos y pluralistas, se puede decir que la evaluación se ha desarrollado, a grandes rasgos, en tres tiempos. En los años sesenta, en la primera generación de políticas de evaluación, el acento se puso en la planificación y la programación de las políticas sociales entonces en boga y, por lo tanto, sobre el papel de la gestión y administración de los programas en cuestión. Para entonces, y sobre todo en Estados Unidos donde algunos servicios públicos alcanzaron un alto grado de sofisticación, se empezaron a afinar técnicas de evaluación que podían distinguirse de las auditorías tradicionales.

En los años setenta, la evaluación comenzó a dirigirse más claramente hacia la satisfacción de dos objetivos. Uno era político: justificar las políticas gubernamentales; el otro era económico: adecuar la administración pública a las crecientes restricciones presupuestales de gobiernos que enfrentarían problemas de escasez de recursos a partir del *shock* petrolero de 1973. El énfasis en la eficiencia se pronunciaría aún más después. En los años ochenta, el tercer momento de las políticas de evaluación estuvo marcado por un interés en la eficiencia, la productividad y los resultados, poniendo el acento en el manejo del presupuesto. Por otro lado, desde la segunda mitad de la década pasada ha sido el poder legislativo el principal impulsor de las evaluaciones, en detrimento de las iniciativas autoevaluadoras de los gobiernos. En ese contexto, los revisores del presupuesto y los auditores de la cuenta

⁴⁷ Todos los recursos, no sólo los financieros: humanos, políticos, militares, legales, sociales, materiales, etcétera.

pública, que rinden informes a las comisiones legislativas, han sido actores centrales para el desarrollo de la evaluación.

La evaluación en Estados Unidos

En este país, la institucionalización de la evaluación es bastante sólida, el grado de complejidad metodológica es elevado y el monto destinado es considerable.⁴⁸ Hoy, el sistema de evaluación constituye una faceta fundamental de las políticas públicas estadounidenses y de los complejos lazos entre el Congreso y la Presidencia.

El origen de las evaluaciones en Estados Unidos se ubica, en los años sesenta, en el marco de los grandes programas sociales (lucha contra la pobreza, etc.) de los presidentes demócratas John F. Kennedy y Lyndon B. Johnson. Dos sucesos coadyuvaron al impulso de su desarrollo. Por una parte, los encargados de las decisiones políticas buscaron controlar mejor los recursos y programas, al dar una racionalidad al proceso administrativo, sobre todo a través del modelo económico integrado en el Sistema de Planeación, Programación y Presupuestación (Planning, Programming & Budgeting System, PPBS). Por otra parte, la aplicación de las ciencias sociales en el manejo de los asuntos del gobierno (en los procesos de planeación y de programación particularmente), así como la utilización de las estrategias de investigación por encuesta y el uso de la estadística también vinieron a reforzar el uso de la evaluación.

En los inicios, el Congreso dependía de las informaciones que el ejecutivo le proporcionaba. Sin embargo, los miembros del poder legislativo recurrieron cada vez más a su Oficina General de Cuentas, OGC (General Accounting Office, GAO), organismo que responde únicamente ante el Congreso. Ello se concretó en 1967 a través de la "ley de oportunidad económica", la cual otorgó al poder legislativo, con la ampliación de las facultades de la OGC, un poder creciente en las diversas áreas de la evaluación de las políticas públicas. A partir de entonces, el trabajo de evaluación se institucionalizó definitivamente en el sistema político administrativo estadounidense.

⁴⁸ Ya a principios de los años ochenta, la Federación norteamericana gastaba más de 300 millones de dólares anuales para evaluar programas y políticas gubernamentales. Este dato tiene el objetivo de dar al lector una idea, aunque haya perdido alguna actualidad, de la inversión en evaluaciones. Véase Garry Brewer y Peter deLeon, *The Foundations of Policy Analysis*, Homewood, Ill., The Dorsey Press, 1983, p. 321.

En la rama ejecutiva una infinidad de unidades distintas llevan a cabo evaluaciones constantemente. Y a pesar de las reducciones considerables de varias de ellas que los gobiernos republicanos de Ronald Reagan y George Bush emprendieron, cada secretaría de Estado cuenta con personal especializado en esta actividad. En la mayor parte de los casos, las evaluaciones son solicitadas por funcionarios de alto rango o por personal administrativo encargado de la operación de cierto programa. El esfuerzo de llevarlas a cabo se manifiesta por el significativo uso de los recursos, la cantidad de personal especializado y la abundante producción de reportes y resultados.

En el poder legislativo cuatro agencias independientes del gobierno que cuentan con una influencia significativa [Oficina General de Cuentas (General Accounting Office), Servicio de Investigación del Congreso (Congressional Research Service), Oficina para el Apoyo Tecnológico (Office of Technology Assessment), Oficina del Presupuesto del Congreso (Congressional Budget Office)] le permiten al Congreso contar con la capacidad de hacer análisis pronosticadores y retrospectivos, contar con diversas opciones para el control de la asignación del presupuesto, efectuar estudios sobre el progreso tecnológico y le ofrecen la posibilidad, en todo momento, de responder a las cuestiones fundamentales relacionadas con la evaluación: ¿qué programas están siendo aplicados, qué servicios suministran, a quiénes benefician, cuál es su grado de eficiencia, qué alternativas están a la mano para cumplir con los objetivos de la política pública? Se cumplen de esta forma, dentro del propio Estado estadounidense, las posibilidades de llamar a cuentas a los responsables de la administración, gracias a las facultades del congreso, con sus fuentes efectivas de poder y contrapesos ante el aparato administrativo del ejecutivo.

La evaluación en la República Federal Alemana

Los orígenes de la evaluación en este país se encuentran en la segunda mitad de los años sesenta, cuando los gobiernos socialdemócratas hicieron esfuerzos por intervenir más activamente en la vida social y económica (auge del Estado benefactor), con lo cual incrementaron de manera significativa los recursos destinados al uso de las ciencias sociales para el control de los efectos de la acción gubernamental. Posteriormente, dos sucesos explican el fortalecimiento de la función de evaluación: 1) que las restricciones económicas obligaron al gobierno

a dirigir los recursos hacia programas más eficientes; 2) que se volvió necesario legitimar algunas políticas públicas que resultaban poco populares y muy onerosas. A los tres motores para el desarrollo de la evaluación en Alemania (análisis de las políticas, uso eficiente de los recursos limitados y necesidad de legitimar las políticas con objetivos sociales), vinieron a añadirse recientemente las presiones del Parlamento alemán —Bundestag—, así como el interés del público (expresado a través de la prensa) sobre ciertos programas en especial.

Por las características del sistema político alemán, cada uno de los ministerios, que conservan una autonomía considerable para formar sus propias estructuras internas, cuenta con grados muy variados de especialización para la evaluación. En el ámbito federal se pueden apreciar tres tipos distintos de institucionalización de la evaluación: 1) en algunos ministerios, donde se han creado unidades especiales para ello; 2) en algunos programas públicos en los que es obligatoria por ley, y 3) la que se ha llevado a cabo en algunos programas sociales experimentales.

Sin embargo, y a pesar de que Alemania cuenta con una estructura fuertemente federalista, la evaluación está muy poco desarrollada en sus territorios —Länder— y a nivel comunitario. Al mismo tiempo, la evaluación en el plano federal está fuertemente desconcentrada, lo cual se debe a la propia estructura administrativa del Estado alemán. El jefe del gobierno —el canciller— tiene funciones que se limitan de manera fundamental a la coordinación, lo cual ofrece una autonomía considerable a los ministerios para la elección, definición, iniciativa y ejecución de las políticas públicas. Cuando se han dado gobiernos de coalición (la mayor parte de los casos desde 1982), inclusive las instancias administrativas regionales han contado con la autonomía suficiente para ejercer un poder de veto frente a las políticas ministeriales.

Por otro lado, el proceso de formulación de las políticas públicas ha hecho difícil que se logre profundizar desde el punto de vista metodológico en las evaluaciones. Las que son externas (las iniciadas por el Bundestag) están diseñadas para cada caso particular y como respuesta adaptada a las necesidades de las políticas coyunturales. Las que son internas, al estar decididas por el Bundestag, siguen por lo regular técnicas cuantitativas, pero más por la premura en las entregas y porque dichas técnicas cuentan con una aceptación generalizada, que por una especial convicción sobre las virtudes de esos métodos. Al mismo tiempo, como una evaluación puede minar la legitimidad, sucede con frecuencia que los resultados se manipulan con fines políticos. Por ello los respon-

sables gubernamentales tienden a mostrar una fuerte reticencia frente a la decisión legislativa de evaluar sus actividades públicas.

Por su parte, la Oficina Federal de Cuentas (Bundesrechnungshof), aunque tiene la obligación de verificar la legalidad de los actos de los organismos gubernamentales y la eficiencia de sus programas, no ha impulsado la evaluación. Se ha adherido así a su tradición auditora al permanecer fuera de los debates políticos y al cuidar celosamente su papel de imparcialidad frente a las acciones gubernamentales.

La evaluación en Canadá

Éste es el ejemplo más próximo a las formas de institucionalización de los estadounidenses. En el rubro de la evaluación, el poder legislativo recuperó los avances que el ejecutivo había perdido. Con un sistema parlamentario, cuyos poderes legislativo y ejecutivo están fuertemente fusionados, Canadá cuenta con un procedimiento de evaluación que se integra totalmente al ciclo de las políticas públicas, el cual fue definido en 1981 en dos documentos que constituyeron desde entonces la guía para las evaluaciones de los programas del gobierno.

No obstante, la evaluación de los programas en el gobierno se desarrolló antes, por iniciativa del Parlamento y del auditor (contador) general, organismo que es independiente del ejecutivo y que cuenta con una influencia considerable. Esta institución había notado desde los años setenta que grandes recursos del Estado se dirigían hacia programas que no gozaban de control alguno. Ello condujo al Parlamento a impulsar las disposiciones legales de 1977 (Ley del Auditor General, Auditor General's Act), con las cuales el auditor podía solicitar la atención y apoyo del propio Parlamento para evaluar las políticas efectuadas de manera insatisfactoria. Para responder a la presión del poder legislativo, el gobierno se vio forzado a crear la Oficina del Contralor General (Office of the Comptroller General, OCG), dotada de una unidad de evaluación de programas (Program Evaluation Branch, PEB). La función de la OCG era asegurar que las disposiciones sobre las evaluaciones de políticas públicas en Canadá se expandieran a todas las agencias y departamentos gubernamentales activos. Adicionalmente, los dos organismos elaboraron dos documentos esenciales que constituyeron las bases para formular la finalidad de la evaluación en la administración pública canadiense.

Dichos lineamientos generales estipulan que todos los departa-

mentos debían desarrollar prácticas de evaluación en sus procesos de acción. De este modo los responsables de los programas públicos, en todos los niveles (directores de empresas públicas, jefes de servicio, viceministros y ministros, etc.), iniciaron políticas de evaluación siguiendo los principios generales. Por otra parte, las actividades de evaluación pueden ser ellas mismas sometidas a una evaluación más formal de la Oficina del Contralor General, del Consejo del Tesoro, de los comités del gabinete o del auditor general. La mayor parte de las evaluaciones son efectuadas por las mismas unidades sectoriales o departamentales (que por lo general cuentan con personal encargado de llevar a cabo esa tarea específica), pero también recurren ocasionalmente a personal externo. Todos los puestos de la administración pública utilizan con regularidad las evaluaciones. En lo que respecta a los destinatarios de sus resultados, existe una ley sobre el acceso a la información de las evaluaciones de políticas públicas que permite a todo el público conocer sus resultados y conclusiones.

Gracias a la presión del poder legislativo, y en particular del auditor general, se institucionalizaron los mecanismos para evaluar las acciones del gobierno, con la posibilidad de que organismos autónomos pidan cuentas a los responsables.

La evaluación en Reino Unido

Desde los inicios del uso de la racionalización y de los métodos cuantitativos para las cuestiones del gobierno, los británicos han puesto el énfasis en el análisis (*ex ante*) de las políticas públicas, más que en la evaluación (*ex post*) de las mismas. Las iniciativas se originaron en los años sesenta, cuando se creó el Programa de Análisis y Revisión, que fue coordinado por el Ministerio del Tesoro. Después de un breve olvido de esa iniciativa, sus actividades fueron reimpulsadas por el enfoque empresarial y eficientista del gobierno conservador de la primera ministra Margaret Thatcher. Por su parte, la Cámara de los Comunes reforzó su interés en la evaluación a partir de la segunda mitad de los años ochenta. Pero las iniciativas en Reino Unido han sido claramente impulsadas por la necesidad de controlar los recursos presupuestales más que por el proyecto de impulsar la función de la evaluación en particular.

Como en los casos anteriores, el desarrollo de la evaluación en este país no puede ser comparado con el de otras naciones sin consi-

derar las particularidades del sistema político británico. De manera especial, deben observarse su carácter parlamentario y las relaciones complejas entre el centro y las regiones, que establecen un buen número de servicios y funciones que son organizados centralmente, pero suministrados por las autoridades locales.

En Reino Unido los ministerios deben responder ante la Cámara de los Comunes, que vigila todas las acciones del ejecutivo. Pero, en realidad, la mayoría parlamentaria cuestiona muy rara vez a un gobierno que pertenece al mismo partido político y que fue constituido a partir de sí misma. Por su parte, los ministerios llevan a cabo la formulación de los programas públicos y utilizan las evaluaciones como instrumentos de consumo interno. Además, y dado que las evaluaciones aportan con frecuencia información importante para la asignación del presupuesto, los ministerios suelen guardar para sí las conclusiones para evitar presiones de la minoría opositora en el Parlamento y los cuestionamientos de la prensa.

Por otra parte, las autoridades locales son responsables en gran medida de sus políticas. Y, por lo tanto, responden en forma individual frente a los electores. Ello explica que la evaluación sea una actividad dispersa. Existe una comunidad modesta de evaluadores dentro y fuera del gobierno que presta sus servicios a demandas ocasionales, pero en realidad no ha habido esfuerzos significativos para que la actividad de evaluación se institucionalice. Dos organismos especiales fueron creados (Unidad de Administración Conjunta, Joint Management Unit; y el Grupo de Administración Financiera, Financial Management Group, que informan al Ministerio del Tesoro), pero tienen un papel esencialmente de orientación, no de control. La función de evaluación tiene poca influencia en la toma de decisiones y su difusión al público es casi inexistente al quedar restringida, casi por lo general, al uso privado del gobierno.

Dentro del Parlamento, el Comité de las Cuentas Públicas (Public Accounts Committee), dotado de la Oficina Nacional de Auditoría (National Audit Office), tiene la facultad de pedir cuentas a los responsables de cualquier programa gubernamental; pero, en realidad, tiene poco personal capacitado y pocos recursos. Desde 1983, dichos organismos pueden inclusive instrumentar la función de evaluación al controlar *ex post* el desempeño económico, la eficiencia y la efectividad de los programas. A pesar de ello, recurren muy poco a esas facultades y, cuando lo hacen, difícilmente pueden catalogarse como evaluaciones rigurosas y con un método definido de análisis.

La evaluación en Francia

La evaluación tenía poca importancia hasta la década de los años ochenta, cuando el interés por ésta se incrementó. La Asamblea Nacional creó la Comisión para el Análisis de Estrategias Científicas y Tecnológicas, y a partir del Noveno Plan Nacional (programa de desarrollo económico, 1983), se integró la idea de evaluar sistemáticamente la instrumentación de políticas. La circunstancia más evidente para explicar el naciente interés en la evaluación la constituyó la cada vez mayor limitación presupuestaria; pero también deben considerarse el creciente involucramiento del gobierno en diversas áreas de la vida social y el aumento en la animadversión por parte del público hacia la intervención del mismo. Sin embargo, hasta ahora ha sido sólo un proyecto: la evaluación no tiene un reconocimiento oficial y por lo tanto no está regulada.⁴⁹ Adicionalmente, la estructura del sistema político administrativo le deja una posición ambigua al evaluador.

En efecto, la política administrativa francesa tiende más hacia la inspección que hacia la evaluación: se verifica con frecuencia el uso del presupuesto y la calidad de la gestión, pero se revisa como un medio de control de los mandos altos a los mandos medios de la administración.

Una de las razones fundamentales es que el poder ejecutivo y el pesado cuerpo de su administración (poco autocrítica y poco inclinada a actuar en forma abierta frente al público) siempre ha predominado sobre el Parlamento: se trata de un sistema de toma de decisiones centralizado y poco inclinado al cambio, a la innovación, a la rectificación y a la publicación de sus resultados y sus cuentas. Por otro lado, difícilmente podría esperarse que la opinión pública o el mismo Parlamento solicitaran una evaluación. Ello ha dependido en realidad de la voluntad de los funcionarios, y la cultura administrativa se ha opuesto a ello: la evaluación está supeditada al proceso de toma de decisiones, sujeta a la administración central. Por otra parte, en las universidades y entre los especialistas de las políticas públicas ha existido poca preocupación o debate sobre el problema de la evaluación y un mercado poco desarrollado y limitado a la propia administración central. Por su parte, el papel de los medios de comunicación también ha sido marginal al expresar poco interés.

⁴⁹ Además, la evaluación ha sido muy poco practicada. De 1971 a 1978, se registraron apenas 36 evaluaciones dignas del nombre. De 1978 a 1983, pudieron contarse más de 150, un número excesivamente modesto.

Al no estar reconocida formalmente como una actividad específica, los funcionarios no asignan en general mayores recursos a la organización de evaluaciones. Además, no se ha generalizado o implementado un sistema metodológico para realizarlas. Los intentos más representativos se han encaminado al campo del análisis cuantitativo y al estudio de los aspectos financieros de un programa, atendiendo a criterios tradicionales de auditoría.

El gobierno (y muy enfáticamente su cuerpo ejecutor, la administración) tiene el papel principal en la demanda de la evaluación; dentro de las agencias públicas son los órganos dedicados a los servicios especializados de investigación los que representan el principal origen de la demanda. Pero la repartición de las evaluaciones es dispareja: algunas áreas de la intervención gubernamental son analizadas de manera constante, mientras que otras igualmente importantes no son controladas o evaluadas jamás. ¿Cuáles caen en el primer grupo y cuáles en el segundo? Tampoco hay regla: la atención especial que pudieran tener algunas políticas en particular depende primordialmente de los intereses convergentes de altos funcionarios de la administración en un momento dado.⁵⁰ Sucede con frecuencia que una política pública es evaluada cuando ha acaparado por algún momento la atención del público a través de los medios de comunicación, por ejemplo.

El hecho de que sea la propia estructura administrativa la que demande el mayor número de evaluaciones tiene varias implicaciones. Por una parte, el estudio se concentra directamente en los objetivos de políticas señalados por la institución particular que demandó la evaluación en lugar de abordar políticas de mayor envergadura cuyo análisis implicaría un aparato conceptual más complejo. Por otra parte, resulta muy raro ver que se lleven a cabo evaluaciones que se interesen en políticas interministeriales. Ello inclusive atañe a la Corte de Cuentas (Cour des Comptes), que en teoría puede iniciar cualquier tipo de investigación que sea de interés para las acciones del gobierno.⁵¹ Sin

⁵⁰ Un caso excepcional fue la iniciativa del Ministerio de la Educación francés que formó el Comité Nacional de Evaluación de las Universidades de Francia para la evaluación de la impartición de educación superior. Estuvo integrado por expertos en ciencias sociales que elaboraron un controvertido informe para el gobierno. Sus conclusiones generales se hicieron públicas, como lo estipulaba el contrato inicial. Al respecto, puede revisarse a Jean Leca *et al.*, "L'évaluation en question", *Revue Française d'Administration Publique*, núm. 66, 1993, pp. 185-226.

⁵¹ La Cour des Comptes es parecida en varios sentidos a la Oficina General de Cuentas (General Accounting Office) de Estados Unidos, pero existen diferencias: aunque depende del Parlamento y está organizada a nivel regional y nacional, tiene también algunas facultades judiciales. Sin embargo, por la conformación de su personal de mayor jerarquía (fruto de la

embargo, ésta fracasa en la explotación de ese poder, pues al estar organizada en departamentos separados y paralelos a la estructura gubernamental, encuentra en su propio seno obstáculos estructurales para desempeñarse como un evaluador versátil.

La evaluación en Dinamarca

El sistema político danés explica el retraso en cuanto a la adopción de la evaluación de las políticas públicas: poca demanda de controles ha resultado en una oferta realmente simbólica. Las decisiones gubernamentales se toman en una estructura formalmente corporativizada que conecta de manera eficaz a las agencias administrativas con los sectores sociales afectados por las políticas. Existen redes adecuadas y bien implantadas para que los sectores sociales establezcan contacto con los funcionarios gubernamentales si, por ejemplo, los programas adoptados tienen efectos nocivos para dichos sectores. Al mismo tiempo, Dinamarca cuenta con mecanismos de revisión que sirven para informar a los funcionarios sobre el desempeño de sus programas.

Los primeros esfuerzos para establecer evaluaciones en las políticas públicas aparecieron a fines de los años setenta. Esta actividad estuvo ligada a las funciones tradicionales de control público sobre la cuenta, a cargo del auditor general, que es independiente del gobierno. Más recientemente, la oficina del auditor general enriqueció sus actividades con análisis que relacionaban la rentabilidad con la calidad, al inspirarse en los modelos de Estados Unidos y Canadá. Por su parte, el gobierno central comenzó a poner en práctica algunas evaluaciones internas. Los casos más significativos lo constituyeron los ministerios del Interior y de Finanzas. El primero buscó en forma explícita mejorar el desempeño de la función pública, mientras que el segundo importó el Sistema de Planeación, Programación y Presupuestación (Planning, Programming & Budgeting System).

Con la llegada del partido conservador al poder, a mediados de los años ochenta, el clima político se caracterizó por un nivel más elevado de conflicto. En ese contexto, el recurso de la evaluación desde algunos comités del Parlamento se hizo más frecuente como instrumento para minar la legitimidad de los opositores políticos. La conse-

tradición imperante en la formación de las élites del Estado que las hace cerradas), su autonomía queda a veces limitada por los propios cuerpos que la integran y que están ligados (por redes de influencia) a otros cuerpos dentro del gobierno.

cuencia fue el crecimiento del interés en cuestiones de rentabilidad, eficiencia y control del gasto público como estrategia de reducción de las dimensiones del Estado y de su intervención en la economía.

La evaluación en los Países Bajos

En el caso de Holanda, el interés por la evaluación como una fase de las políticas públicas es reciente y su origen está en el Parlamento. Ello se logró al ofrecer a la Cámara General de Cuentas (Algemeen Rekenkamer) autonomía legal frente al gobierno y al Parlamento, y al convertirla en una entidad subsidiaria del presupuesto del poder legislativo. En principio, su función es la verificación de la legalidad de las acciones gubernamentales, pero cada vez más se ha visto en la necesidad de evaluar también la eficiencia y de informar al público al respecto.

Anteriormente, el Ministerio de Finanzas había intentado promover, sin éxito, el análisis de los programas gubernamentales. En los años ochenta, el gobierno conservador tomó la iniciativa de institucionalizar la evaluación con el proceso anual del presupuesto. Mientras tanto, tal y como sucede en Reino Unido, los esfuerzos de hacer la administración pública más eficaz estuvieron ligados al interés político de revertir una larga tradición de intervencionismo del Estado en la economía. Así como sucedió en otros países, la evaluación fue más bien identificada con la noción de control al presupuesto y a la legalidad en el ejercicio del gasto.

La introducción de la evaluación tuvo lugar a principios de los años setenta con la instauración de un comité interministerial para el análisis de los objetivos políticos que interesaban a más de un departamento a la vez: su función estuvo ligada más a la necesidad de coordinar acciones dentro de la administración que a las actividades propias de la evaluación. El intento de aplicar métodos analíticos a las políticas públicas se topó con diversas resistencias en los círculos administrativos holandeses. Sin embargo, cuando la crisis económica se profundizó, a fines de la década de los años setenta, y ante la necesidad concomitante de reducir el déficit presupuestario, el gobierno se vio motivado a financiar, de manera marginal, la investigación en estrategias de evaluación. La implementación en 1981 de un procedimiento llamado *Reconsideraciones* se orientó en este sentido. Esta iniciativa estaba destinada a reducir el déficit fiscal y a dictar claras instrucciones obligatorias para todos los departamentos administrativos.

La evaluación en la Confederación de Suiza

El sistema administrativo suizo se caracteriza por ser una organización fuertemente federalista, cuyos fundamentos son los siguientes: las administraciones locales se encargan de casi la totalidad de las acciones gubernamentales que interesan a la comunidad, respetando con escurpulosidad el rigor financiero que siempre ha caracterizado a la gestión de los asuntos públicos, aunque existe el principio de *subsidiaridad*,⁵² para el apoyo ocasional de las acciones públicas. Este sistema implica que la actuación de la autoridad y la ejecución de sus programas está muy cerca de los ciudadanos.

En casos extraordinarios, la Confederación (gobierno central) puede asistir a los cantones para resolver algunas cuestiones de política social o inversiones que son demasiado onerosas. Pero la mayoría de los asuntos públicos y casi la totalidad de los servicios se suministran por autoridades locales, gracias a un papel muy activo de la sociedad civil, a través de un sistema de asociaciones y organizaciones muy eficaz.

El *consenso*, principio clave para la regulación social y la convivencia política en Suiza, implica generalmente que las decisiones son tomadas por la autoridad después de amplias consultas con los sectores interesados (afectados). El papel del Parlamento (la Asamblea Federal) es el de ser la agencia de supervisión superior, para lo que cuenta con varios comités de control a su disposición, los cuales, en su mayoría, tienen una capacidad de análisis retrospectivo de las acciones públicas, lo que significa que el Parlamento puede señalar los errores después de una evaluación *ex post*, pero difícilmente puede hacer que una política pública local sea modificada o cancelada a lo largo de su aplicación, pues en realidad no cuenta con los recursos suficientes para asumir una vigilancia más activa que comprometa recursos y personal activo a lo largo de la instrumentación.

Por otra parte, la Confederación Suiza no tiene una Corte de Cuentas o Contaduría Mayor que dependa de su poder legislativo. El control sobre los recursos y los gastos públicos se garantiza a través de una división del Departamento Federal de Finanzas que es relativamente autónoma, en tanto que el control parlamentario sobre las finanzas se ejerce a través de un Comité de Finanzas para cada sector y de un

⁵² Dicho principio obliga a la autoridad federal a intervenir para robustecer al gobierno local o completar su acción, como último recurso.

Comité Rector: la Oficina del Delegado de Finanzas. Sin embargo, dichos órganos se centran casi exclusivamente en el control presupuestario: no logran en realidad explicar las variaciones entre objetivos de la política pública y los efectos concretos, ni permiten un cuestionamiento demasiado exhaustivo de los gastos.

La evaluación en Noruega

El gobierno noruego ha centrado su atención, de manera constante, en la planificación desde los años sesenta. De nueva cuenta, fue la crisis de fines de los años setenta la que obligó al gobierno a mirar hacia la función de la evaluación. A partir de 1985, el Parlamento (Storting) hizo obligatoria la entrega de informes sobre actividades gubernamentales, entre los que se consideraron criterios ligados a la evaluación. Sin embargo, el origen del interés parlamentario estuvo en las limitaciones económicas y en la necesidad de afinar el control sobre las actividades presupuestales gubernamentales.

En el poder legislativo, la institución que efectúa las evaluaciones es la Oficina del Auditor General, que depende del Parlamento y que vigila las actividades de las entidades individuales del gobierno federal en todas las áreas que tienen una relación con la cuenta pública. Pero sus esfuerzos interesan principalmente el manejo financiero. Por parte del gobierno ha existido una creciente preocupación por llevar a cabo evaluaciones, pero los alcances han sido modestos.

Aspectos determinantes para el desarrollo de la evaluación

A partir de la revisión anterior, vamos a identificar y a discutir seis rasgos del surgimiento de la evaluación institucional de las políticas públicas. En primer lugar, casi siempre ineludible, está el factor presupuestario. Los casos que surgieron en la primera ola de evaluaciones en los años sesenta fueron sustentados en el crecimiento económico que permitió a los gobiernos llevar a cabo sustantivas políticas de intervención en el ámbito social. En ese momento surgieron los primeros intentos por integrar la evaluación a las diversas fases de la política pública, desde su programación hasta su terminación. En la década siguiente, la de los años setenta, al llegar los tiempos de la escasez económica, las situaciones presupuestales críticas trajeron el desarro-

llo de otra forma de evaluación, esta vez ligada al proceso de programación y vigilancia del gasto, así como a la búsqueda de eficiencia y rentabilidad en la función pública.

El segundo rasgo sería la condicionante política. De manera tal vez coincidente con el marco económico, el contexto político pareció influir también en el surgimiento y multiplicación de la función de evaluación. En los años sesenta, las evaluaciones de políticas públicas aparecieron por primera vez cuando llegaron al poder grupos políticos reformistas (los demócratas en Estados Unidos, los socialdemócratas en la República Federal Alemana, los liberales en Canadá), lo cual estableció un ambiente político más propicio para que las ciencias sociales se utilizaran para la toma de decisiones gubernamentales. En sentido contrario, pero equivalente, los gobiernos conservadores de los años ochenta, de frente a las dificultades económicas, asumieron varias reformas que implicaron el retiro de la economía por parte del Estado, y dieron prioridad a la rentabilidad y a la eficiencia de las políticas públicas.

Sin embargo, algunos autores han considerado que el argumento de la coyuntura política constituye una explicación demasiado simplista e insuficiente para explicar la integración de la evaluación de programas por los distintos gobiernos que han recurrido a ellas.⁵³ Para ellos, y éste sería nuestro tercer rasgo característico, lo que resulta primordial es, entre otras consideraciones estructurales, la relación entre el poder ejecutivo y el poder legislativo. En los sistemas tradicionalmente parlamentarios, el gobierno cuenta con un control bastante considerable sobre la mayoría parlamentaria (la cual, se sobreentiende, pertenece al mismo partido político), mientras que la oposición, minoritaria, está en una situación desventajosa para controlar a las comisiones que se encargan de los controles y evaluaciones sobre los programas gubernamentales. Es el caso de Canadá y Reino Unido, Dinamarca, Noruega y los Países Bajos. La tendencia general es, por lo tanto, a que el desarrollo de la evaluación sea menos importante que en los sistemas presidencialistas (como en Estados Unidos).

En estos últimos, la oposición política en el Congreso tiende a or-

⁵³ Entre ellos Hans-Ulrich Derlien, quien no cree que el color del partido en el poder influya sobre el desarrollo de la evaluación gubernamental. En su crítica a P. Wagner y H. Wollman ("Fluctuations in the Development of Evaluation Research: Do 'Regime Shifts' Matter?", *International Social Science Journal*, 1986b, pp. 601-617), Derlien se inclina a hacer énfasis en las cuestiones de tipo institucional como los factores determinantes en el desarrollo de las evaluaciones de políticas públicas. Véase H. U. Derlien, en R. C. Rist, *op. cit.*, pp. 147-175.

ganizarse mejor para frenar o controlar al ejecutivo en vista de su dominio casi total sobre el gobierno. Cuando el sistema lo permite, el poder legislativo puede constituirse en una columna considerable para construir un sistema institucionalizado de evaluaciones al gobierno. Como ha sucedido en Estados Unidos, esto es posible inclusive si los esfuerzos de evaluación se han originado en el seno del gobierno. El interés de la rama legislativa se explica no sólo por la necesidad de controlar al ejecutivo a través de supervisiones a las evaluaciones que este último se hace sino para conservar un poder de información equivalente al que tiene el propio gobierno, para así preservar su independencia.

Es claro que dicho marco explicativo debe ser aceptado con reservas. Por una parte, sucede con frecuencia que en algunos sistemas tradicionalmente parlamentarios, el supuesto control que el gobierno ejerce sobre la mayoría legislativa se ve cuestionado de manera rigurosa en algunos asuntos que dividen a esa mayoría (pensemos solamente en los *euroescépticos* británicos o los *québécois* en Ottawa). Por otra parte, algunos sistemas fuertemente presidencialistas no han visto desarrollarse la evaluación, como en el caso de Estados Unidos. Así tenemos los casos del Canadá parlamentarista y de la Francia presidencialista: en el primer caso, la evaluación está bien implantada en las instituciones de los poderes legislativo y ejecutivo, mientras que en Francia, la evaluación no hace sino comenzar a formar parte del proceso de sus políticas públicas. Para interpretar con justicia esos dos casos, deben también considerarse condiciones como la tradición política, la cultura en la gestión y administración pública y la formación profesional de las élites gobernantes, tanto en el poder ejecutivo como en el legislativo.

Otro factor constitucional (o legal) muy importante —el cuarto— es el papel y la filiación de la Oficina General de las Cuentas Públicas. En la mayoría de los países considerados, los organismos de control al presupuesto son entidades que dependen del Parlamento y responden al poder legislativo (los casos de Estados Unidos, Canadá, Reino Unido, Noruega, Holanda y Dinamarca). En los casos de Francia y Alemania, aunque las contadurías responden ante la Asamblea, se inclinan con más facilidad (por disposiciones legales o por tradición política) hacia el poder ejecutivo; en el caso de Suiza tiene predominancia la autoridad local. Pero puede decirse que el nivel de independencia de la Oficina de la Cuenta Pública frente al gobierno y la administración explica la tendencia del Parlamento a utilizar fuentes de información alternativas a los institutos dominados por el ejecutivo.

La quinta característica determinante de los diferentes procesos de evaluación es la disponibilidad de especialistas y científicos sociales, y la relación de su gremio con el gobierno.⁵⁴ Al ser necesario un nivel de especialización muy considerable para llevar a cabo evaluaciones, el desarrollo de las mismas está decisivamente ligado a los instrumentos producidos en el sector académico y en los centros de investigación social.

Finalmente, en sexto lugar, importa la existencia (o la ausencia) de organismos que, en el seno del gobierno, tengan funciones equivalentes a la evaluación. Es el caso de las oficinas públicas dotadas de sistemas de estadística y de información, que sean relativamente autónomas y eficientes. En este rubro, importa también la existencia de grupos de asesores gubernamentales que cuenten con una infraestructura técnica elevada, la existencia de empresas dedicadas a la asesoría y la consultoría (que realicen evaluaciones de políticas públicas) y, desde luego, la existencia de un periodismo dedicado a la investigación que pueda ocasionalmente sustituir a las evaluaciones gubernamentales (sucede en Estados Unidos, Canadá y Suecia, por ejemplo).

Resumen esquemático de los rasgos característicos de la evaluación en nueve países

Nota explicativa

Este resumen intenta condensar la información sobre la evaluación en los nueve países. Es importante indicar que subraya las líneas generales para cada país, con lo que se corre el riesgo evidente de incurrir en generalizaciones (sobre todo en el recuadro de características).

1) La *definición* de evaluación que proponemos respeta el uso más frecuente que de ella se hace en cada cultura político-administrativa. En algunos casos, es proporcionada explícitamente por los autores que participaron en el trabajo editado por R. C. Rist.

2) Dentro de las *condiciones estructurales* consideramos sobre todo las características de los sistemas políticos y la realidad constitucional (incluyendo el grado de independencia del auditor de la contaduría del Congreso o la Oficina General de Cuentas Públicas respecto al gobierno).

3) En cuanto al *papel de las evaluaciones en la toma de decisiones*, dis-

tinguimos entre las evaluaciones destinadas al análisis de la totalidad de un programa y las que buscan incrementar la eficiencia de las políticas públicas en condiciones de limitaciones presupuestales.

4) El *nivel de institucionalización* se refiere a la existencia de organismos especializados en la función de evaluación, así como a la de leyes que obliguen a los gobiernos a evaluar sus actuaciones y programas de acción.

5) El *proceso de evaluación* se basa en las cuatro preguntas que propone R. C. Rist:

- ¿Qué autoridad solicita las evaluaciones?
- ¿Qué unidad o grupo las instrumenta?
- ¿De dónde provienen los recursos necesarios?
- ¿Quién utiliza las informaciones generadas?

6) Las *características* consideradas son las siguientes (aquí atenderemos simplemente a una tendencia general) (véanse las definiciones en el Anexo):

- Internas o externas.
- Sumarias o formativas.
- Criterios endógenos o exógenos.
- Métodos formales (oficiales) o adaptaciones al caso.

7) El *entorno social* considera la disponibilidad para la evaluación de investigadores en ciencias sociales, la formación profesional de los encargados de la decisión y de la política, las dimensiones del país y el desarrollo de alternativas y de sustituciones.

8) Y la *difusión* de las conclusiones de la evaluación, que puede ser abierta o limitada.

Estados Unidos

❖ Definición

La evaluación permite identificar:

- 1) Los programas públicos que se están aplicando.
- 2) Los servicios que dichos programas están efectivamente proporcionando.
- 3) Los beneficiarios verdaderos de esos servicios y de esos programas.
- 4) La eficacia, eficiencia y rentabilidad en la puesta en práctica de esos programas.
- 5) El costo social de la política pública evaluada (por lo menos, lo que se tiene que sacrificar para cumplir con ella).

❖ Condiciones estructurales

Son óptimas. Es un sistema presidencialista, con políticas públicas bastante centralizadas, donde la fracción legislativa de oposición puede

⁵⁴ Este argumento lo debemos a P. Wagner y H. Wollman, *idem*.

fácilmente cuestionar al gobierno y a la administración. El Congreso puede también reforzar las capacidades de control sobre el ejecutivo. La Oficina General de Cuentas (que responde ante el poder legislativo), está obligada legalmente (Ley de Oportunidad Económica, 1967) a evaluar a la administración pública.

- ❖ **Papel de la evaluación en las decisiones**
Influencia muy considerable en todos los niveles de gobierno, con énfasis en la eficiencia de los programas. Desde un punto de vista político, las evaluaciones pueden ser utilizadas para legitimizar acciones gubernamentales (por la administración), o para atacarlas (por los adversarios en el Congreso).
- ❖ **Nivel de institucionalización**
Muy alto. El Congreso se apoya sobre cuatro instituciones de análisis e información: la Oficina General de Cuentas, el Servicio de Investigación del Congreso, la Oficina para el Apoyo Tecnológico y la Oficina de Presupuesto del Congreso. Por su parte, el gobierno cuenta con personal especializado y numeroso para las funciones de evaluación. La institución evaluadora más cercana al presidente es la Oficina de Administración y Presupuesto.
- ❖ **Proceso institucional de la evaluación**
 - Origen de la solicitud: gobierno y Congreso.
 - Evaluadores: miembros de la administración, Congreso, instituciones especializadas no gubernamentales, centros de investigación universitaria, empresas consultoras o de asesoría, etcétera.
 - Financiamiento: gobierno, Congreso e instituciones privadas.
 - Utilización: Congreso, gobierno, empresas privadas, opinión pública.
- ❖ **Características**
Toda la gama de posibilidades: internas o externas; sumarias o formativas; prospectivas o retrospectivas; con criterios endógenos o exógenos; con teoría de la evaluación o con asesoría de un experto, etcétera.
- ❖ **Entorno social**
Existe una disponibilidad muy importante de investigadores en ciencias sociales y especialistas, pues hay una demanda general de los mismos. Se utilizan constantemente las ciencias aplicadas o cuantitativas para las políticas públicas (desde los años sesenta). Los responsables de las políticas en muchas ocasiones son sujetos altamente especializados (economistas y científicos sociales, muy frecuentemente). Y existen importantes alternativas de instituciones de evaluación, como algunos organismos privados (*think-tanks*), los centros académicos y la prensa.
- ❖ **Difusión**
Totalmente abierta al público. Sólo se prohíbe oficialmente en las áreas que comprometen la "seguridad nacional".

Alemania

- ❖ **Definición**
La evaluación es un mecanismo instrumental de revisión con el cual los administradores y los encargados de la decisión pueden juzgar el desempeño de un programa. Por lo general, el gobierno alemán utiliza métodos científicos para la búsqueda de la eficiencia de los programas públicos y el análisis de sus efectos.
- ❖ **Condiciones estructurales**
Son relativamente favorables. Es un sistema político bastante descentralizado, donde el Parlamento puede cuestionar al jefe del gobierno que, en realidad, coordina a los ministerios, integrados a partir de las coaliciones políticas. Los ministerios y los territorios cuentan con bastante autonomía para elaborar sus políticas públicas. Pero llevan a cabo pocas evaluaciones. Por su parte, el Parlamento no cuenta con el personal o las unidades necesarias para evaluar a la administración, y la Oficina General de Cuentas, dependiente del legislativo considerablemente autónomo, no recurre a la función de la evaluación, se concentra en funciones de auditoría.
- ❖ **Papel de la evaluación en las decisiones**
Considerable en el nivel sectorial (por áreas de actuación de cada ministerio). Para algunos programas legislativos, el Parlamento solicita evaluaciones a las instancias gubernamentales correspondientes, con especial interés en el problema de la eficiencia. Es frecuente que algunas evaluaciones experimentales se vean politizadas en el debate parlamentario.
- ❖ **Nivel de institucionalización**
Fragmentado. El Ministerio de Finanzas tiene atribuciones legales (1973) para pedir evaluaciones a algunos programas sectoriales aplicados por otros ministerios, pero recurre a éstas esporádicamente. Algunos ministerios tienen unidades especiales de evaluación.
- ❖ **Proceso institucional de la evaluación**
Origen de la solicitud: en primer lugar, el Congreso; después, los ministerios.
 - Evaluadores: los ministerios y diversas instancias ejecutoras.
 - Financiamiento: el gobierno federal.
 - Utilización: Parlamento, ministerios y cancillería (jefatura del gobierno).
- ❖ **Características**
Evaluaciones internas y externas. Sumarias (Parlamento, gobierno) y formativas (políticas experimentales, llevadas a cabo por el gobierno). Criterios endógenos y teoría evaluativa.

- ❖ Entorno social
Disponibilidad de investigadores en ciencias sociales. Utilización de las ciencias sociales aplicadas y cuantitativas en las políticas públicas. Recientemente, economistas y científicos sociales se han integrado al grupo de decisión en el gobierno.
- ❖ Difusión
Teóricamente, limitada al Parlamento y al gobierno; pero, en realidad, existe un acceso cada vez más considerable para el público.

Canadá

- ❖ Definición
El objetivo esencial de la evaluación es proporcionar a los usuarios (funcionarios) y al público una información suficiente y oportuna para mejorar la distribución de los recursos y facilitar la toma de decisiones que perfeccionan los programas.
- ❖ Condiciones estructurales
Favorables. Es un sistema parlamentario en el cual los poderes ejecutivo y legislativo están relativamente integrados. Las evaluaciones tienen un soporte legal (Ley del Auditor General, 1977). El auditor general, que responde ante el Parlamento, ha servido de impulsor de la evaluación a programas públicos.
- ❖ Papel de la evaluación en las decisiones
Influencia muy considerable. Todas las instancias gubernamentales están obligadas a llevar a cabo evaluaciones, para la revisión tanto de la eficiencia como de la rentabilidad.
- ❖ Nivel de institucionalización
Muy alto. Existe la Unidad de Evaluación de Programas, que fue organizada por la oficina gubernamental de contraloría, y que integra las funciones de evaluación (Oficina del Contralor General, creada en 1978). La evaluación hace parte suya el ciclo general de la gestión para la administración pública: existe una guía de principios y de instrumentación de evaluaciones para los funcionarios.
- ❖ Proceso institucional de la evaluación
 - Origen de la solicitud: ministros, altos funcionarios, directores de paraestatales, etcétera.
 - Evaluadores: agencias gubernamentales, la Oficina de la Contraloría General que controla el proceso, empresas de consultoría, grupos de investigación universitaria, etcétera.
 - Financiamiento: se integra al presupuesto de las oficinas gubernamentales.
 - Utilización: todas las instancias del gobierno.

- ❖ Características
Son generalmente internas (en las oficinas de gobierno). Sumarias. Con criterios endógenos, siguiendo dos documentos guía o de instrucción: Guía de la Función de la Evaluación Programática y Principios para la Evaluación de Programas por Departamentos y Agencias Federales de 1981. Se ha desarrollado, por lo tanto, una teoría evaluativa.
- ❖ Entorno social
Existe una alta disponibilidad de investigadores en ciencias sociales. Utilización de las ciencias sociales aplicadas en las políticas públicas. Economistas y científicos sociales se integran cada vez más a la élite gobernante y a la toma de decisiones.
- ❖ Difusión
Totalmente abierta.

Reino Unido

- ❖ Definición
La evaluación está ligada al proceso de control de presupuesto así como a la apreciación de la responsabilidad del funcionario público (su desempeño legal y administrativo, o *accountability*).
- ❖ Condiciones estructurales
Poco favorables: la mayoría parlamentaria partidista ha mostrado poco interés en cuestionar a un gabinete que se forma a partir de ella misma. La Oficina Nacional de Auditoría, que se dirige a la Cámara de los Comunes, puede pedir cuentas a los responsables de cualquier programa gubernamental, pero cuenta con poco personal capacitado y pocos recursos.
- ❖ Papel de la evaluación en las decisiones
La evaluación tiene poca influencia. Los ministerios guardan para circulación interna los resultados de sus autoevaluaciones. Se ha dado importancia a la eficiencia de las políticas públicas.
- ❖ Nivel de institucionalización
Fragmentado en cuanto a sus métodos y participantes, disperso en cuanto a los organismos involucrados.
- ❖ Proceso institucional de la evaluación
 - Origen de la solicitud: ministerios y direcciones generales.
 - Evaluadores: las direcciones o departamentos encargados de la política o programa.
 - Financiamiento: se integra al presupuesto de los ministerios.
 - Utilización: limitada a los ministerios e instancias oficiales del gobierno.

- ❖ **Características**
Internas, sumarias, criterios endógenos (Unidad de Administración Conjunta, 1985, y Grupo de Administración Financiera, 1987), con metodología formal (teoría de la evaluación).
- ❖ **Entorno social**
Existe poca disponibilidad de investigadores sociales dispuestos a participar en puestos públicos (existe un sistema de carrera civil muy rígido para funcionarios). Se recurre poco a las ciencias sociales aplicadas para las políticas públicas. Formación clásica de los funcionarios encargados de las decisiones.
- ❖ **Difusión**
La mayor parte de las evaluaciones constituyen informaciones administrativas clasificadas (Ley de la Información Pública y Secreta) y fácilmente se convierten en documentos con acceso restringido al público.

Francia

- ❖ **Definición**
La evaluación es toda actividad que consiste en registrar los efectos de la política pública, analizarlos de acuerdo con un criterio preestablecido, elaborar un informe y presentarlo al comanditario que solicitó la evaluación.
- ❖ **Condiciones estructurales**
Poco favorables: el presidente, el primer ministro y el pesado cuerpo de la administración siempre han predominado sobre el Parlamento, al constituir un sistema de toma de decisiones centralizado y poco inclinado hacia la publicación de sus resultados. Marcado énfasis en la noción de control sobre los criterios de evaluación, organizado a partir de instancias gubernamentales. La Corte de Cuentas, que responde ante la Asamblea Nacional, se restringe a funciones de control sobre uso de recursos y legalidad de las actuaciones del gobierno.
- ❖ **Papel de la evaluación en las decisiones**
Influencia moderada e irregular. Algunas áreas de la actuación pública son analizadas constantemente, mientras que otras, que son de primera importancia, no lo son. Predomina un criterio discrecional de los altos responsables políticos en cuanto a la elección de los programas sujetos a evaluación.
- ❖ **Nivel de institucionalización**
Modesto: dentro de la cultura de la administración, la evaluación no está reconocida como una actividad integral de las políticas públicas. La Asamblea Nacional no ha impulsado las instancias legislativas que podrían crecer en este sentido.

- ❖ **Proceso institucional de la evaluación**
 - Origen de la solicitud: la estructura administrativa del Estado: funcionarios y encargados de programas.
 - Evaluadores: las diversas instancias gubernamentales responsables de los programas que requieran de la evaluación.
 - Financiamiento: generalmente deben integrarse al presupuesto del programa evaluado.
 - Utilización: limitada a los órganos de la administración.
- ❖ **Características**
Internas, sumarias, con criterios endógenos y metodología formal.
- ❖ **Entorno social**
Las universidades e investigadores sociales han mostrado poco interés en el problema de la evaluación; hay un mercado poco desarrollado y limitado a la propia administración central.
- ❖ **Difusión**
Poco interés en los medios de comunicación y fuerte control del Estado (de la administración) sobre la información que compete su desempeño.

Dinamarca

- ❖ **Definición**
La evaluación es el análisis sistemático de la aplicación y los resultados de las políticas públicas.
- ❖ **Condiciones estructurales**
Poco favorables: sistema parlamentario que busca tradicionalmente evitar los conflictos entre el poder legislativo y el gobierno (y la representación partidaria del gobierno en el Parlamento no tiene interés en confrontar al ejecutivo). Se trata de un sistema corporativo que ha relacionado eficazmente a las agencias administrativas con los sectores sociales afectados. El auditor general es independiente del gobierno, pero se restringe generalmente a controles legales y presupuestales.
- ❖ **Papel de la evaluación en las decisiones**
Marginal, con énfasis en el control de la eficiencia de las políticas públicas.
- ❖ **Nivel de institucionalización**
Modesto. Se recurre esporádicamente a la evaluación por el auditor general; con más frecuencia, el Ministerio de Finanzas lleva a cabo algunas.
- ❖ **Proceso institucional de la evaluación**
 - Origen de la solicitud: auditor general o ministerios de Finanzas o de Desarrollo.
 - Evaluadores: las instancias gubernamentales.
 - Financiamiento: presupuestos de los programas del gobierno.

- Utilización: Parlamento y gobierno.
- ❖ Características
 - Internas, sumarias, criterios endógenos y metodología formal.
- ❖ Entorno social
 - Existen mecanismos de revisión y de información alternativos para los responsables de las políticas (los que toman las decisiones). La formación profesional de los altos funcionarios es de legalistas, por lo que están poco familiarizados con la evaluación. La aplicación de los métodos cuantitativos en las políticas públicas es muy reciente y existe una pobre disponibilidad de investigadores en ciencias sociales para los asuntos de la administración pública.
- ❖ Difusión
 - Limitada, con un interés modesto por parte del público.

Países Bajos

- ❖ Definición
 - La evaluación es parte del análisis de pronóstico y retrospectivo de las políticas públicas, a fin de mejorar la calidad y el desempeño de éstas, así como los procedimientos administrativos.
- ❖ Condiciones estructurales
 - Favorables. El Parlamento puede cuestionar al gobierno, que tiene, no obstante, cierto margen de maniobra ejecutiva. La Segunda Cámara del Parlamento, encargada de aprobar el presupuesto, tiene facultades para vigilar también su aplicación. La Cámara General de Cuentas es bastante independiente, tanto del Parlamento como del gobierno, pero su papel se restringe a la verificación de la legalidad de las cuentas públicas.
- ❖ Papel de la evaluación en las decisiones
 - Modesto, pero creciente en lo que respecta a la eficacia administrativa del gobierno. Se ha puesto el énfasis en la eficiencia y la disminución del déficit presupuestal.
- ❖ Nivel de institucionalización
 - Aún limitado, pero debe anotarse la creación, en 1981, del Procedimiento de Reconsideraciones, operación que ha consistido en la revisión de políticas para reducir los gastos, con instrucciones obligatorias para todas las instancias públicas interesadas. En algunas ocasiones, dichas instrucciones incluyen evaluaciones que se fundan en criterios racionales (y no exclusivamente financieros).
- ❖ Proceso institucional de la evaluación
 - Origen de la solicitud: Parlamento, Cámara General de Cuentas, Ministerio de Finanzas.
 - Evaluadores: instancias gubernamentales.

- Financiamiento: presupuesto central.
- Utilización: Parlamento y gobierno.
- ❖ Características
 - Internas, sumarias, criterios endógenos y metodología formal.
- ❖ Entorno social
 - Existe una modesta disponibilidad de investigadores en ciencias sociales dedicados al tema. Pero hay un sustituto: el comité de consejeros en torno al gobierno central. En general, la formación profesional de los funcionarios es de abogados, por lo que hay poca familiarización con la evaluación.
- ❖ Difusión
 - Abierta a la opinión, a través de informes que la Cámara General de Cuentas publica con regularidad.

Suiza

- ❖ Definición
 - La evaluación reúne una serie de mecanismos de control que varias agencias gubernamentales proporcionan en forma explícita.
- ❖ Condiciones estructurales
 - Poco favorables. Sistema fuertemente federalizado en el que las políticas públicas son instrumentadas localmente (la autoridad federal interviene sólo en caso extremo, bajo el principio de la *subsidiaridad*). La mayor parte de los asuntos públicos se resuelven a través de un control muy activo de las organizaciones y sectores sociales afectados (la sociedad civil). Ni el Parlamento ni la Confederación (poder ejecutivo con poderes limitados para establecer programas nacionales) cuentan con los recursos técnicos, humanos y materiales suficientes para vigilar activamente a los cantones (autoridades locales).
- ❖ Nivel de institucionalización
 - Muy modesta, aunque existen intentos incipientes. No obstante, las reglas y procedimientos que hay son imprecisos y no están formalizados. En 1987 se estableció un organismo de investigadores por iniciativa del Departamento de Justicia (rango ministerial); en 1989, se creó un programa nacional de investigación para optimizar los recursos públicos.
- ❖ Proceso institucional de la evaluación
 - Origen de la solicitud: la sociedad civil.
 - Evaluadores: los propios funcionarios o comités extraparlamentarios.
 - Financiamiento: la Confederación y la Asamblea Nacional.
 - Utilización: el Parlamento y la sociedad civil.
- ❖ Características
 - Externas, sumarias, criterios exógenos y metodología formal.

- ❖ Entorno social
La estructura política fuertemente descentralizada inhibe el desarrollo de criterios generales y sistemáticos. Por otra parte, la sociedad tiene otras formas de control, a través de su participación activa en la implementación de gran número de políticas públicas.
- ❖ Difusión
Completamente abierta al público.

Noruega

- ❖ Definición
La evaluación busca garantizar que, a lo largo del proceso de la política pública, haya correspondencia entre los objetivos y el trabajo gubernamental para lograrlos.
- ❖ Condiciones estructurales
Poco favorables. Es un sistema parlamentario en el cual la mayoría, al pertenecer al partido que gobierna, se caracteriza por su prudencia en los juicios sobre el desempeño del ejecutivo. Al mismo tiempo, y dada la conformación del Congreso, la minoría opositora no tiene los medios. Por su parte, la Oficina del Auditor General responde al legislativo, con atribuciones principalmente legales y presupuestales.
- ❖ Papel de la evaluación en las decisiones
Su función es secundaria; se centra en cuestiones de eficiencia, de control al presupuesto y de productividad para algunos programas públicos.
- ❖ Nivel de institucionalización
Modesto y reciente. En algunas instancias gubernamentales se lleva a cabo investigación retrospectiva.
- ❖ Proceso institucional de la evaluación
 - Origen de la solicitud: Parlamento y gobierno.
 - Evaluadores: instancias gubernamentales, instituciones independientes de investigación, empresas de asesoría, grupos universitarios, etcétera.
 - Financiamiento: presupuesto central.
 - Utilización: Parlamento y gobierno.
- ❖ Características
Internas y externas; sumarias, con criterios exógenos. Teoría de evaluación, aunque se recurre con frecuencia a la opinión de expertos (el modelo del *experto* de Monnier, ya aludido en la nota al pie número 38).
- ❖ Entorno social
Existe una disponibilidad creciente de investigadores especializados, pero hay una tradición de cortes administrativas compuestas por fun-

cionarios, principalmente abogados, que sustituyen las evaluaciones con base en revisiones con múltiples criterios.

- ❖ Difusión
Está abierta en algunos rubros que conciernen al presupuesto, pero el gobierno tiene la capacidad de limitarla en algunos casos, cuando resulta conveniente hacerlo por razones de "interés general".

Conclusiones

Una queja frecuente de los ciudadanos es que los funcionarios no tienen representatividad: ¿quién eligió al director general de crédito público de la Secretaría de Economía? Sin embargo, una decisión de ese personaje puede afectar a miles de gobernados con todo y sus bolsillos. Los funcionarios trabajan encerrados en sus despachos y toman decisiones que afectan a los gobernados. Cuentan con la distancia protectora del escritorio y se aíslan tras los obstáculos físicos y simbólicos de las oficinas públicas; se esconden detrás del anonimato de las instituciones del Estado; y con la complicidad de los demás funcionarios de arriba y de abajo de la jerarquía, los funcionarios que deciden, que actúan o se inmovilizan con indiferencia, gozan de una fuerte dosis de impunidad.

Por ello, es frecuente que a los funcionarios se les reclame porque su nombramiento depende de sus superiores (el jefe del gobierno o los miembros del gabinete) sin pasar por el escrutinio público, por lo que no pueden ser fácilmente cuestionados. Con dificultad pueden ser llamados a rendir cuentas acerca de sus actuaciones. En realidad, los ciudadanos tienen pocas posibilidades y recursos legales para hacerlos responder en forma efectiva por sus acciones y pedirles explicaciones amplias y razones convincentes. Desde el punto de vista de la sociedad, esto puede generar la triste paradoja de que mientras alimenta al Estado con los impuestos que paga, tiene que soportar al mismo tiempo autoridades públicas a las que no les puede pedir cuentas ni grados de responsabilidad. Y es que el recurso ciudadano que sirve por excelencia para premiar o castigar a los gobernantes, el voto, no es atribuible a la mayoría de las autoridades públicas. Y cuando un funcionario es elegible, el periodo que existe entre una elección y la siguiente es demasiado prolongado, lo cual deja manos libres al gobernante por largo rato.

Una de las virtudes de integrar institucionalmente la evaluación en el ciclo de las políticas públicas es que permite cuestionar las acti-

vidades de esos funcionarios. Si la evaluación de un programa gubernamental es obligatoria y se lleva a cabo por un tercero, dentro de un marco de criterios de evaluación claros, por lo menos se puede discernir acerca de la conveniencia de continuar con el programa o suspenderlo, y se puede calificar o descalificar la actuación de las autoridades que lo decidieron y ejecutaron. De esta manera, la evaluación obliga a los funcionarios a mejorar su desempeño público en sus actuaciones y en la utilización de los recursos escasos. Sin embargo, para que la evaluación funcione como un medio efectivo de control de las acciones públicas, deben cumplirse tres condiciones: 1) que el proceso de evaluación esté implantado en forma institucional y que sea vigilado por lo menos por una instancia del Estado distinta al organismo ejecutor; 2) que las conclusiones de la evaluación del desempeño de la unidad gubernamental queden al alcance de los ciudadanos, y 3) que el marco legal existente garantice los mecanismos de verificación, control y juicio para que, en caso necesario, se sancione al responsable de los excesos y las fallas.

Las razones que hacen deseable la evaluación institucionalizada en los programas del gobierno son también las que, en parte, han obstaculizado su desarrollo: por lo general, las administraciones y autoridades se rehúsan a someter sus decisiones y acciones a la luz evaluadora de criterios que pueden venir a cuestionarlas. Por ello, inclusive en varios países democráticos la evaluación de políticas públicas ha tenido un desarrollo limitado. Por otra parte, debe decirse que la evaluación no es la única forma de plantear en forma efectiva y práctica el problema de la *accountability*. La evaluación es una forma de pedir cuentas al gobierno, pero no es la única. En Suiza, lo hemos visto, las evaluaciones gubernamentales están poco desarrolladas, pero los niveles de control de los actos de autoridad por parte de la sociedad agrupada en asociaciones y en organizaciones civiles son muy elevados y eficaces.

Dice José Antonio Crespo que un elemento esencial de la democracia es su capacidad institucional para pedir cuentas —legal y políticamente— a los gobernantes por la vía pacífica.⁵⁵ Esta revisión de la evaluación en nueve países ha dejado claro por lo menos lo siguiente: el desarrollo de la evaluación como forma de control de la acción gubernamental ha demandado disposición y esfuerzo dentro del propio

⁵⁵ Lo ha escrito en varias ocasiones. Por ejemplo, en el diario mexicano *Reforma* del 6 de marzo de 1995.

gobierno y, de manera determinante, dentro del poder legislativo. La actitud del público en general frente a las evaluaciones, su solicitud o indiferencia, también ha sido un motor fundamental de su desarrollo. Y para que dichos esfuerzos den frutos, ha sido necesario establecer un marco de lucha política institucionalizada y reglas del juego legales y claras para la competencia por el dominio de los aparatos y recursos del Estado. En ocasiones la oposición ha impulsado desde el Congreso el desarrollo de la evaluación; a veces la oposición recién llegada al gobierno ha condenado las acciones del pasado o justificado sus acciones del presente con el recurso de la evaluación institucional. Siempre, el pluralismo y la democracia, si bien no han sido factores suficientes para la evaluación de las políticas públicas, han sido condiciones necesarias.

Anexo. Definiciones útiles

- *Evaluación descriptiva*: registra los cambios generados por una política pública.
- *Evaluación clínica*: explica los efectos de la política pública a partir de sus objetivos declarados y utilizando los criterios y valores oficiales.
- *Evaluación normativa*: explica los efectos de la política con los criterios que el analista de la evaluación elige.
- *Evaluación experimental*: busca las relaciones estables de causalidad entre la política pública y sus efectos, a través de experimentos y ensayos con un grupo tratado que se compara con un grupo de referencia.
- *Evaluación pluralista*: adopta los puntos de vista de varias categorías de sujetos, grupos o personas interesadas por la política pública para establecer una interpretación múltiple, dinámica y consensual de sus efectos.
- *Objetivos*: metas, fines, circunstancias o situaciones que se busca alcanzar, ya sea porque se consideren deseables por sí mismos o porque resulten prácticos.
- *Criterios*: reglas, referencias o escalas que sirven para graduar, interpretar y calificar lo que se observa. Los criterios de la evaluación se llevan a efecto —se operacionalizan— a través de métodos de análisis, observación, registro o medición.

- **Técnicas (o métodos):** procedimientos, medios e instrumentos que permiten poner en práctica la evaluación. Se derivan, generalmente en forma coherente, de los criterios.
- **Valores:** graduación, puntuación, interpretación o calificación de la política pública con base en las referencias de valoración o criterios. A partir de éstos se determinan las características distintivas, se miden los méritos y se calculan los desaciertos de un programa de actuación gubernamental. El valor puede referirse a la importancia, la calidad, la sustancia, la cantidad, la estimación, la utilidad, la justicia, la eficacia, etc., de la política pública.
- **Normas:** preceptos de conducta que gobiernan y dirigen el comportamiento al asegurar la regulación de la vida de los individuos y los grupos dentro de la sociedad. Cuando varias personas aceptan una determinada jerarquización de valores (lo bueno y lo malo, lo superior y lo inferior, etc.) surgen las *normas* (aunque en una organización social, pueden ser definidas impositivamente, por ley).
- **Criterio de eficiencia:** logro de objetivos al menor costo posible.
- **Criterio de eficacia:** logro de objetivos concretos.
- **Criterio de igualdad:** mismo tratamiento a todos los sectores sociales, aunque éstos se encuentren en situación desigual.
- **Criterio de equidad:** busca eliminar la injusticia al favorecer explícitamente a los sectores marginados; se preocupa en particular por las situaciones de desigualdad social, para lo cual integra consideraciones distributivas.
- **Criterio de pertinencia o adecuación:** se refiere a la capacidad de responder de manera satisfactoria a las necesidades o demandas sociales.
- **Criterios endógenos:** definidos a partir de las metas propuestas por los diseñadores de la política en cuestión, por lo que son congruentes con los criterios internos de la institución involucrada.
- **Criterios exógenos:** provienen de consideraciones externas a las que norman y justifican la política pública evaluada.
- **Evaluación interna:** elaborada por instancias que son dependientes de la organización que aplica la política.
- **Evaluación externa:** llevada a cabo por un organismo que no

- está integrado y que es autónomo respecto del cuerpo público responsable de la política por evaluar.
- **Evaluación sumaria:** consiste en el análisis de las consecuencias generales de la política en cuestión al poner el énfasis en los efectos finales de la política, después de considerar el estado de cosas anterior (v.g., un solo examen final para evaluar un programa).
 - **Evaluación formativa:** permite establecer observaciones y recomendaciones tácticas a los responsables de la política (los que decidieron ponerla en práctica), a fin de corregir desviaciones y optimizar el desempeño del programa (v.g., varias inspecciones parciales cuyos resultados se promedian para evaluar un programa).
 - **Métodos o procedimientos formales:** sucede cuando las unidades o instituciones dedicadas a la evaluación recurren a una técnica generalizada (como una teoría reconocida), y definida en forma legal u oficial.
 - **Métodos ad hoc (convicción íntima):** sucede cuando las instituciones o unidades evaluadoras adaptan sus técnicas al caso específico en el que se ocupan, sin que haya un procedimiento riguroso y general que cumplirse.