

Una comparación internacional en la evaluación de las políticas públicas *

Hans-Ulrich Derlien **

I. Introducción

Hay tres funciones que se atribuyen a la evaluación de las políticas: de información, de reasignación y de legitimación (que se explicarán a continuación). La *tesis de este artículo* es que, desde los años 60 a los 90, el énfasis político sobre la evaluación ha pasado desde la función de información a la de reasignación.

Este cambio de posición se puede apreciar desde *dos momentos distintos* en los que la evaluación se institucionalizó como una herramienta de gestión en los gobiernos nacionales. Detrás de este cambio hay razones políticas y económicas específicas. Además, junto con el cambio de motivos hay un cambio de actores dominantes (sección II).

Los años 90 se han caracterizado por *la persistencia y el declive* de la evaluación. A pesar de la ulterior profesionalización, e incluso popularización, de la evaluación en Europa, la situación actual en muchos países muestra el predominio de la función de reasignación (vinculada al movimiento de la *New Public Management*). Al mismo tiempo, los evaluadores tienden a volverse auditores con un énfasis metodológico en la medición de los resultados (*output*), mientras que, en un principio, los auditores utilizaron técnicas de evaluación en la mayoría de los países avanzados. Esto afectaría a la calidad de la evaluación tanto como a la evaluación de la calidad (sección III).

1. Función de información

La función de información predominó durante la fase de despegue del movimiento de evaluación en los años 60. Desde

el punto de vista organizativo estaba situada en los departamentos gubernamentales responsables del proyecto y la ejecución de las políticas de reforma.

Como consecuencia de ello, los gestores de los programas se hacían las siguientes preguntas:

- ¿Funcionan las políticas?, es decir, ¿qué efecto producen?, en particular, ¿se han conseguido los objetivos programados (los efectos deseados)?, ¿qué consecuencias no previstas, positivas o negativas aparecen?
- ¿Cómo se pueden mejorar estas políticas, una vez que los estudios de evaluación han puesto de manifiesto los fallos?

Esta función de información de la evaluación se agrupa alrededor de nuevas políticas puesto que la predicción de los resultados y, por tanto, los análisis previos son difíciles de llevar a cabo. Porque las políticas de reforma:

- a) se emprenden principalmente en las áreas de políticas «blandas» como salud, educación, bienestar, donde no se dispone de pruebas científicas para evaluar relaciones de causa-efecto en el área en cuestión;
- b) además, a menudo, estas políticas eran innovadoras y no suponían simplemente pequeñas variaciones a partir de vías políticas trilladas, lo que las exponía a determinados riesgos de fracaso.

2. Función de asignación

La función de evaluación predominante en los años 80 durante la segunda etapa de institucionalizaciones estaba destina-

da claramente a ayudar a realizar una asignación racional del presupuesto. Ahora, los actores principales no son ya los administradores de los programas, sino las oficinas auditoras, los Ministerios de Hacienda y las unidades centrales, a quienes les compete la elaboración global del presupuesto y su estructura interna.

Las cuestiones básicas que se plantean ahora son:

- ¿Qué programas pueden recortarse o contenerse a partir de los resultados negativos de la evaluación?
- ¿Cuáles son las consecuencias del «retroceso de las fronteras del Estado», es decir, la privatización de determinadas actividades públicas?
- ¿Cómo podemos sacarle «más rendimiento al dinero» mediante la reorganización de los programas individuales y de las actividades públicas en su conjunto?

Las cuestiones inducidas de los motivos de asignación tienden a poner más énfasis en la cantidad que en la calidad.

3. Función de legitimación

Contrariamente a lo que ocurre con las funciones anteriormente mencionadas, la legitimación no es claramente atribuible como motivo a las fases específicas de desarrollo de la institucionalización de la de evaluación de programas; siempre juega un papel importante puesto que la evaluación de programas tiene lugar en un contexto político. Se cree que las pruebas científicas justifican las decisiones políticas, bien sea para mejorar, disminuir o suprimir programas. Sin embargo, como consecuencia de la «ley de hierro de los fallos metodológicos de la evaluación», es más difícil de probar que los programas y las políticas son efectivos, que lo contrario; por tanto, es más fácil deslegitimar políticamente las políticas públicas que legitimarlas realmente bajo condiciones de escasez fiscal.

Depende mucho de si los evaluadores se convierten en auditores o si los auditores actúan incorporando la evaluación. Durante los años 90 en Europa, algunos sistemas nacionales de evaluación estuvieron fuertemente influenciados por el auditor dominante: *evaluación imperativa de los fondos estructurales de la Unión Europea*.

II. Las dos etapas de la institucionalización

Durante los años 90 pueden observarse tres tendencias en el proceso de institucionalización de la evaluación de los programas en los países de la OCDE:

- En los años 80 aparece un segundo movimiento que se ocupa de cuestiones de *output* y, en menor medida, del impacto: en el Reino Unido durante el mandato de la Sra. Thatcher, en los Países Bajos y en Noruega. Incluso, la mayoría de los países que institucionalizaron la evaluación a finales de los años 60, experimentaron un segundo cambio de actividades de evaluación, como en Suecia.
- Mientras que el primer movimiento de evaluación de los años 60 y 70 (EE.UU., Canadá, Suecia y Alemania) estuvo muy unido al proceso de planificación y programación y, por tanto, al papel de los administradores de los programas, los intentos a finales de los años 70 y 80 se orientan más bien a la reconsideración, a nivel político, de la justificación de las políticas o del proceso presupuestario. En ambos casos el papel de los auditores externos se convirtió en importante.
- Consecuentemente, el enfoque de la evaluación ha variado en los años 80, pasando de ser, originariamente, una operación interna del gobierno a ser un asunto de los Parlamentos. Unido a esto, los actores también han variado. Ahora el Parlamento es el centro clave de actividad y, además, las oficinas auditoras, que antes informaban al Parlamento, han sido incluidas en el sistema.

1. Explicación de la emergencia de la evaluación

¿Qué factores determinaron la maduración de los sistemas de evaluación y qué constelaciones influyeron en los modelos de desarrollo de los distintos países?

1.1. La situación fiscal

Las distintas pautas de desarrollo de los países que comenzaron más temprano —EE.EE., Canadá y la República Federal Alemana— que los distingue de la mayor parte de los otros países que les siguieron a finales de los años 70 o más tarde, podrían estar relacionadas, entre otros factores, con las condiciones de las economías nacionales y, por tanto, con la situación fiscal de los gobiernos respectivos.

El primer grupo de países impulsó la evaluación de políticas durante los años 60; es decir, países con una situación económica próspera y con presupuestos en crecimiento que permitían a los gobiernos embarcarse en costosos programas de intervención social, incluyendo educación y salud. En este contexto surgieron sistemas formales de planificación que, o estaban limitados a una planificación económica a medio plazo (República Federal Alemana) o, intentaban integrar el presupuesto con la programa-

ción (EE.EE., Suecia y Canadá). En cualquier caso, la evaluación o se consideraba, lógicamente, como parte de estos sistemas de planificación o se requería para la información de los programas de intervención, cuyos resultados podrían predecirse con suficiente grado de certeza. Las evaluaciones, entonces, se utilizaban principalmente por los administradores de programas para llevar a cabo los programas ya existentes y otros nuevos. Sólo en ocasiones fueron necesarias las evaluaciones para proteger estos programas de los oponentes políticos, para probar su efectividad.

El impacto de la por completo diferente situación económica de finales de los años 70 sobre la evaluación fue casi el mismo: las reducidas economías y la necesidad de gestionar la escasez, produjo tensiones en los gobiernos y les llevó a desarrollar técnicas para gestionar los recortes (SCHICK, 1988). De esta forma, la situación económica relativamente desfavorable y la crítica situación fiscal promovieron que la evaluación se hiciera por vía del proceso presupuestario: las iniciativas de gestión financiera (Gran Bretaña, Canadá y Noruega, Países Bajos) o bien hicieron uso de esfuerzos de evaluación anteriores (Canadá y Gran Bretaña) y dieron lugar a una segunda ola de esfuerzos de evaluación, o el nuevo gerencialismo introdujo el concepto de evaluación para racionalizar la distribución de los recursos dentro del presupuesto (Noruega y Países Bajos). Los que soportaron este segundo movimiento de evaluación no son los administradores de los programas en los departamentos gubernamentales, sino más bien los ministros de Hacienda y las oficinas auditoras como guardianes tradicionales del presupuesto. Por consiguiente, la perspectiva y la función de evaluación cambiaron ligeramente: en lugar de poner en práctica programas, el énfasis se ponía más bien en la reducción de los programas ineficaces para, de esta forma, recortar el presupuesto nacional.

1.2. Constelación política

Por supuesto, no es la situación económica *per se* la que ocasiona cambios en la dirección política, sino que, en último extremo, son los políticos y los partidos en el gobierno los que cambian el curso de la nave del Estado.

En los prósperos años 60, los países de la primera ola de la evaluación estaban gobernados por partidos reformistas (Social Demócratas en la República Federal Alemana y en Suecia, Demócratas en EE.UU, Liberales en Canadá), que acometieron en las diversas políticas de reforma y tenían una clara afinidad para el empleo de los métodos de las ciencias sociales para la evaluación de las políticas. Otro caso pertinente podría ser el de la presidencia socialista en Francia (1981) que puso énfasis en la evaluación en 1982. Los gobiernos más *Conservadores*, que se enfrentaron, a menudo, a una situación económica menos prós-

pera, hicieron uso de los instrumentos de evaluación existentes, no sólo para reelaborar el presupuesto sino, en ocasiones, para intentar reducir los programas de reforma de los gobiernos predecesores, que desde el principio estaban mal vistos. Esto no significa que si los partidos políticos intervencionistas hubiesen estado en una posición de gobierno, no hubiesen tenido que responder de manera similar a la nueva situación económica.

No es sorprendente que exista también un consenso neo-conservador internacional, consistente en creer que el Estado debería retirarse de la sociedad o, al menos, no debería expansionarse más, y que muchos programas habrían ido más allá de sus objetivos. En conexión con las políticas de privatización, desregulación y desburocratización, la fragmentación del sistema de planificación, que había servido para desarrollar los programas de reforma, fue un método para cambiar el curso de la nave del Estado. El otro método fue *mejorar el track presupuestario tradicional mediante la importación de técnicas de auditoría* del sector privado —técnicas que tenían que tratar con los resultados de los programas y (menos con el impacto), como la unidad de análisis y, por tanto, se parecían a los métodos de evaluación. A fin de cuentas, las modernas técnicas de auditoría encajaban con los procedimientos tradicionales de las auditorías de las oficinas auditoras—.

1.3. Características constitucionales

Una explicación de las distintas vías de desarrollo seguidas por los diferentes sistemas nacionales que se concentran únicamente en la correlación entre una economía en expansión/programas intervencionistas/gobiernos social-demócratas o una economía en recesión/política de austeridad/gobierno conservador sería un modelo demasiado sencillo. Por ejemplo, esto no explicaría el cambio que comenzó en los EE.UU. bajo la Administración Republicana de Nixon ya en 1970, tampoco explicaría por qué casi nada cambió después de que un gobierno conservador llegase al poder en Bonn en 1982. Una variable adicional de importancia primordial parece ser la *relación constitucional entre los poderes ejecutivo y legislativo del gobierno*.

En sistemas parlamentarios de *gobierno de partido* la teoría clásica de la división de poderes difícilmente es cierta en la práctica; en este caso el gabinete ejerce un fuerte control sobre la mayoría parlamentaria o comparte intereses y un consenso programático de las facciones mayoritarias; sólo los grupos de la oposición tienen una motivación importante para controlar al ejecutivo, aunque normalmente no están en situación de imponerle controles. Esta breve descripción histórica esclarece el papel transmisor de los Parlamentos en los sistemas de gobierno de partido (Canadá, Noruega, Gran Bretaña). Sin embargo, en

todos los casos las evaluaciones reales se llevaron a cabo por el poder ejecutivo.

La situación podría ser completamente diferente en sistemas de *gobierno presidencial* como en EE.UU y Francia donde pueden crearse mayorías parlamentarias en contra del Presidente e imponer medidas sobre una Administración para acrecentar su capacidad de control o independencia de información. No es de extrañar que el poderoso Congreso de los Estados Unidos reuniese un enorme equipo en el mismo Congreso e hiciese uso de esta contra-burocracia para controlar a los administradores de programas en la rama ejecutiva. En tal situación el poder legislativo aparece como un pilar esencial en el sistema de evaluación, incluso cuando las iniciativas de evaluación se originan en la Administración. El interés del legislativo puede ayudar a mantener en forma apropiada las capacidades de evaluación del poder ejecutivo, incluso si la Administración no se muestra demasiado entusiasta con ello, puesto que necesita de estos equipos de análisis para mantener el equilibrio de poder de información *vis à vis* con el Congreso.

Otro factor constitucional que podría ser de importancia a este respecto es la *asociación del Tribunal de Cuentas* o la unidad auditora. En la mayoría de los países estas oficinas se consideran como unidades de apoyo al poder legislativo e informan al Parlamento (EE.UU, Canadá, Noruega, Países Bajos y Dinamarca). La *General Accounting Office* en la República Federal Alemana ha estado históricamente vinculada al ejecutivo y sólo desde 1969 orientó ligeramente su función para servir también al Parlamento y al reducido comité auditor parlamentario. De forma similar, la *Cour des Comptes* francesa, aunque también informa al Parlamento, tiene una tendencia hacia el ejecutivo debido a la formación profesional de sus miembros (NIOCHE/POINSARD, 1985). En Suecia, con su tradición de pequeños ministerios y agencias muy independientes, la institución auditora dominante, la Oficina Nacional Auditora (*National Audit Bureau*) es, exclusivamente, un instrumento del gobierno central. Sus informes se envían, tanto a la agencia auditora como al gobierno cuando se trata de asuntos importantes. Los auditores parlamentarios cuentan, en este caso, con un pequeño grupo de asesores para llevar a cabo escrutinios *ad hoc*.

Cuanto más independiente del ejecutivo es la unidad central auditora y vinculada al Parlamento, más tentados están los Parlamentos para usar estas instituciones como una base de poder y como una fuente de contra-asesoramiento en pie de igualdad, incluso, con un gobierno de partido. En el caso americano esta disposición, por supuesto, no haría sino fortalecer las iniciativas de evaluación del Congreso.

También vale la pena considerar si la *Constitución federal o unitaria* de un Estado tiene efecto sobre el grado en que se aplica la evaluación de los programas. En nuestras consideraciones podemos incluir también en esta categoría a Estados descentraliza-

dos que establecieron *gobiernos regionales* como España, Italia y Francia. ¿Tienen los centros del Estado en sistemas federativos o regionales mayor necesidad de controlar las actuaciones de los gobiernos regionales? Sería razonable suponer que existe tal necesidad de información en los casos en que los programas se llevan a cabo por oficinas de los gobiernos regionales (Alemania, Canadá, Suiza), pero que existiese menos necesidad cuando el gobierno central ejecuta los programas a través de sus propias oficinas (EE.UU). Así, en Canadá y en la República Federal Alemana, donde se han puesto en marcha programas conjuntos, federales y estatales, sobre una base de costes compartidos, se institucionalizaron las evaluaciones para asegurar que los programas alcanzan los niveles nacionales. Es aleccionador que el gobierno federal suizo en 1987 institucionalizara un *think tank* (centro de investigación) para explorar las posibilidades de evaluación y después creara un gran fondo para evaluación. En otros países, la política regional europea (fondos de cohesión) desencadenó las necesidades de evaluación del Estado central.

El control se lleva a cabo con más facilidad en los Estados unitarios donde el Estado central puede requerir estadísticas de las oficinas regionales, mientras que en un sistema federal con autonomía administrativa de los miembros, no es posible, en ocasiones, al menos políticamente, solicitar información detallada de los Estados regionales, a menos que la obtención esté específicamente institucionalizada.

2. Dispositivos de evaluación en el Gobierno

En este epígrafe se hará un intento para complementar la perspectiva genérica mediante un informe sistemático de las adaptaciones de evaluación desarrolladas en varios países. Se puede estudiar mejor cómo se integra la evaluación en la maquinaria del gobierno, al analizar *la distribución de las iniciativas de evaluación*.

A menos que la realización de los estudios de evaluación esté institucionalizada, éstos suelen tener lugar de manera aleatoria. Además, el número de estímulos potenciales para estos estudios va desde la crítica pública a la necesidad genuina que tiene el administrador de programas de saber cómo funcionó el programa en la práctica. Las siguientes observaciones estarán limitadas a las iniciativas que tienen lugar porque la función de evaluación esté institucionalizada de algún modo en el sistema político-administrativo.

2.1. Obligación legal y requerimiento parlamentario para evaluar

Debido al papel destacado que han tenido los Parlamentos para poner en marcha evaluaciones ordinarias en la mayoría de

los países, no es frecuente que el mandato para llevar a cabo una evaluación se establezca en una ley específica que autorice el programa respectivo. En estos casos, el poder ejecutivo está obligado a elaborar un *informe al Parlamento* después de un cierto período de tiempo o a intervalos regulares; puesto que la ley no siempre especifica los contenidos del informe, la tarea se ha interpretado en general como una evaluación. Por ejemplo, la mayor parte de los 200 informes ordinarios del ejecutivo en Alemania (de los cuales no todos son evaluaciones) dirigidos al Parlamento tienen su origen en peticiones parlamentarias permanentes. La práctica para institucionalizar la evaluación en la legislación que autoriza se ha hecho en los Estados Unidos mediante el sistema de convertirla en *sunset legislation* como política general del Congreso. Con esta estrategia política, los programas que no han probado su efectividad en una evaluación expiran automáticamente después de un determinado número de años.

Puesto que los estudios de evaluación casi siempre conllevan la reunión del análisis de grandes cantidades de datos, los Parlamentos, generalmente, piden a la burocracia del gobierno que haga la investigación, y con frecuencia la investigación real queda delegada posteriormente a especialistas externos. Existe, sin embargo, una excepción a la regla: en los casos en que el Parlamento por sí mismo mantenga un gran equipo de investigación y oficinas de apoyo, las iniciativas serán, a menudo, dirigidas hacia estos especialistas del legislativo, como el caso de EE.UU. En estas circunstancias hay posibles instancias que lleven a cabo evaluaciones paralelas en el ejecutivo, así como en las oficinas de apoyo del legislativo.

2.2. *Iniciativas de las Instituciones supremas de auditoría*

No está claro hasta qué punto las oficinas auditoras, que han sido de una importancia máxima para la institucionalización de la evaluación, inician y llevan a cabo las evaluaciones por ellas mismas. Indudablemente, cuanto más promueva la oficina auditora las actividades de evaluación general del ejecutivo, más *aconsejará* al gobierno y a los ministerios, más tarde, para evaluar medidas específicas, a menos que el departamento quiera solicitar crítica pública del Tribunal de Cuentas. Por otro lado, en los procesos reales de auditoría los estudios de evaluación son llevados a cabo posiblemente por los propios auditores, dependiendo de su experiencia para comprometerse en esta tarea.

En contraste con la *General Accounting Office* de EE.UU., la mayoría de las oficinas auditoras en Europa, con su tradición de algunos siglos, pueden estar menos a favor de adoptar la nueva terminología y, por tanto, parece que han estado bastante inactivas en este campo hasta fecha reciente. En la práctica,

no obstante, su trabajo puede que no sea tan diferente (aunque los números son definitivamente más pequeños) del de la *General Accounting Office* de los EE.UU. Sin embargo, suele ser poco frecuente que se evalúen programas en su totalidad por las oficinas auditoras europeas excepto, posiblemente, para gastos en programas de equipos de defensa.

2.3. *Iniciativas del Ministerio de Hacienda*

Otra fuente de iniciativas pueden ser los Ministerios de Hacienda o las oficinas equivalentes como la OMB en Washington. Por supuesto, los informes de las oficinas auditoras se retroalimentan en el proceso presupuestario, es decir, los análisis posteriores tendrán un impacto en las asignaciones presupuestarias e, incluso, en las propuestas presupuestarias del Ministerio de Hacienda. La relación entre evaluación y creación del presupuesto cambiará si el Ministerio de Hacienda está en posición de solicitar evaluaciones de los administradores de programas en los distintos departamentos. Como se ha indicado anteriormente, ésta es la situación en la República Federal Alemana. El grado de relación entre la creación del presupuesto y la evaluación depende de la frecuencia y regularidad con la que se solicitan estudios; es una cuestión de negociación entre los Ministerios de Hacienda y los relacionados con los programas, y es posible que se sacrifiquen algunas evaluaciones para minimizar conflictos. Una solución *estructural* para tratar este problema puede encontrarse en la instauración de una Agencia para el Análisis de Políticas en el Ministerio de Hacienda holandés, mientras que el sistema danés de rotación anual de evaluaciones desde las políticas en otra área constituye un dispositivo *procedimental* para unir la evaluación de manera más segura a la creación del presupuesto.

2.4. *Iniciativas políticas centrales*

El que se tome la iniciativa para institucionalizar los procedimientos de evaluación en el gobierno no implica que el jefe del gobierno o la Oficina del Presidente como actores políticos centrales no inicien también estudios *específicos* de evaluación.

Lo que es posible que hagan los actores políticos centrales es tomar la iniciativa para «evaluar» las políticas más que los programas, es decir, tratan con áreas de políticas o problemas más amplios que, frecuentemente, involucran distintas jurisdicciones departamentales. Están interesados en análisis comprensivos relacionados con estas cuestiones. Cuanto más amplio y menos específico sea el problema que se considere, más probable es que su tratamiento tenga lugar en comisiones y comités. El razonamiento inverso sugiere que las evaluaciones de programas específicos podrían hacerse por especialistas de la casa.

Ejemplos de un enfoque global y un evaluador especialista central se pueden encontrar en Gran Bretaña: el antiguo *Central Policy Review Staff* era una institución que informaba directamente al Primer Ministro, revisaba problemas de política amplios y desarrollaba nuevas opciones. La *Efficiency Unit* siguió esta línea de trabajo después de 1979; su intento para hacer que la maquinaria del gobierno fuese más eficaz implicaba, sin embargo, que se realizaran menos estudios sustantivos y más estudios organizativos de reforma dirigidos, entre otras cosas, a la autoevaluación de los departamentos.

Desde 1981, el *Gabinete holandés* decide sobre los temas que deben evaluarse por los grupos de estudio en los departamentos dentro del Procedimiento Anual de Reconsideración. También debe mencionarse aquí la OMB en Washington ya que lleva a cabo funciones presupuestarias centrales y es una oficina del gobierno central. Obviamente, la OMB es un organismo central de iniciativas dado que solicita evaluaciones de los departamentos.

En conclusión, podría mantenerse que en la mayor parte de los países no se toman iniciativas para evaluar programas específicos y proyectos. En los años 80, las excepciones a la regla eran:

- a) La OMB de EE.UU. cuyas iniciativas derivan de su función presupuestaria más que de su función como órgano político,
- b) la *British Efficiency Unit* que inicia evaluaciones en la medida en que sean parte de la política para mejorar las operaciones de gestión, y
- c) el Gabinete holandés que toma decisiones específicas de evaluación.

En general se puede manifestar que el núcleo del gobierno, aparte de poner énfasis en la necesidad de evaluar e institucionalizar la función, genera *en el mejor de los casos evaluaciones amplias de las políticas*, que por sí mismas no cumplen los requerimientos formales de una evaluación completa. Estas mediciones y evaluaciones, como se llaman con frecuencia, se preparan después por unidades de personal del gobierno central o por consejeros y comités.

2.5. Iniciativas de los departamentos

Como se mencionó más arriba, los departamentos tienden a evitar la autoevaluación; por esta razón, los Parlamentos y sus agencias de apoyo, así como los gabinetes, institucionalizaron o propagaron evaluaciones de los departamentos en el pasado. Por tanto, muchas de lo que puede parecer a primera vista como iniciativas departamentales están de hecho *inducidas externamente*. Un ejemplo de esto son las agencias centrales sue-

cas que son responsables ante el gobierno de la evaluación de sus programas. No obstante, hay iniciativas genuinas de los departamentos. Los esfuerzos de evaluación con respecto a los programas de intervención de los años 60 parecen haberse originado en los departamentos respectivos, tanto más cuanto mayor es la afinidad entre los administradores de programas y las investigaciones externas relevantes que trataban con el problema para el cual se había diseñado el nuevo programa.

Desde una perspectiva amplia de gobierno, estas iniciativas descentralizadas se agrupaban en ciertas áreas de políticas (educación, empleo, sanidad, ayuda al desarrollo) y no involucraban a la mayoría de los departamentos. Para estabilizar las iniciativas autónomas intradepartamentales o para iniciar actividades de evaluación en los departamentos que permanecían inactivos hasta el momento, se crearon *unidades de evaluación*. Estas unidades comprenden desde aquellas que planifican y diseñan estudios y escriben informes sin realizar necesariamente la investigación, hasta aquellas que se ocupan de todos los aspectos de las evaluaciones.

- a) Existen equipos de personal (206) en todos los departamentos norteamericanos no relacionados con la defensa.
- b) Se establecieron unidades de evaluación en algunos ministerios alemanes (Ayuda al Desarrollo, Tráfico, Oficina de Prensa, Agricultura, Tecnología).
- c) En Dinamarca existen estas unidades en los departamentos de Ayuda al Desarrollo, Trabajo y el Consejo de Bienestar (ALBAEK/WINTER, 1988).
- d) El Ministerio de Cooperación para el Desarrollo en Noruega creó una división de evaluación en 1983.
- e) El que los Subsecretarios en Canadá fuesen nombrados responsables de las evaluaciones del departamento se puede considerar como una alternativa funcional para los equipos de personal.
- f) En Suecia existen grupos de expertos en la mayoría de los ministerios que están implicados en actividades de evaluación.
- g) En los Países Bajos los estudios en los departamentos se preparan bajo el Procedimiento de Reconsideración en cuyo trabajo participan funcionarios del Ministerio de Hacienda.

De esta manera, en la mayor parte de los sistemas políticos nos encontramos una *multiplicidad de grupos de interés (stake holders)*. Esta fragmentación concuerda con el panorama que habíamos expuesto con respecto al *proceso* de institucionalización. Aparte de institucionalizaciones externas e iniciativas recurrentes, el mismo ejecutivo toma iniciativas para evaluar sus propias actividades (por decirlo de manera cuidadosa). En cuanto a las unidades centrales de gestión, las iniciativas pueden origi-

narse en el Ministerio de Hacienda y convertirse en parte de la rutina presupuestaria. Los actores políticos centrales, sin embargo, tienden a limitarse a fortalecer la función de evaluación del departamento.

III. Desarrollos en los años 90

La teoría de las dos olas de evaluación estableció, entre otros factores, un paso entre las economías en expansión y las economías en recesión, por un lado, y la naturaleza cambiante de la evaluación, por otro. Actualmente, la situación es mucho *más ambivalente*.

1. El declive de la evaluación: un escenario pesimista

Con la redefinición de las actividades del Estado y su alcance, nadie está ya interesado en la mejora del programa; de lo que se trata es de la reducción de programas. Esto se relaciona con la segunda etapa.

Como consecuencia de esto, internacionalmente existe una tendencia a desconectar la política de las políticas; basta echar una ojeada a la retórica de las políticas: Tony Blair predica la compasión y al mismo tiempo está recortando los programas sociales.

Además, la *New Public Management* podría haber contribuido al declive de las medidas de impacto puesto que los políticos pretenden fiarse más y más de las medidas del mercado.

Las necesidades de la gestión fiscal, en particular en Alemania desde 1990, restan atención al *output* y al impacto de la puesta en práctica de las políticas. Se convierten en fundamentales las cuestiones sistemáticas relacionadas con el mantenimiento del sistema de seguridad social. Las prioridades están en el *input* mientras que los *outputs* ya no tienen interés a no ser que se conviertan en un problema para continuar asegurando los *inputs*.

El análisis de impacto es menos atractivo cuanto menos conscientes son los políticos y el público de las circunstancias distributivas de los impuestos y del gasto.

Además, la evaluación de programas como retroalimentación sistemática en el desarrollo de las políticas *entra en declive en periodos de crisis*; es, obviamente, un mecanismo para periodos de desarrollo normal de políticas y de aprendizaje (en el sentido de Thomas KUHN). Para la gestión fiscal de crisis, entre otros resultados negativos de la evaluación, cualquier argumento sirve para reducir programas. Al mismo tiempo, los indicadores de impacto se clasifican por las medidas de *output* como en la *New Public Management*: sus «indicadores de funcionamiento»

son en todo caso medidas de *output*, que dejan fuera consideraciones sobre los efectos distributivos.

Esta tendencia está acompañada de una política de desregulación que significa, en esencia, abandonar los mercados a sí mismos y permitir más externalización en la sociedad.

2. Persistencia de los sistemas de evaluación

No obstante, las dos configuraciones que formaron los movimientos de los años 60 y 80 de institucionalización de la evaluación de las políticas en la maquinaria del gobierno nacional, no cambiaron en todos los países durante la última década de este siglo. A pesar de esto, sí hubo cambios.

- En parte, los nuevos países adoptaron la evaluación de las políticas o la evaluación de las políticas había madurado hasta tal punto que se reconoció internacionalmente y se informaba de ello (Israel, Austria, Irlanda, Unión Europea). Siguiendo los motivos de la segunda etapa y apoyada por la *New Public Management*, Australia y, posiblemente, Nueva Zelanda empezaron a considerar la evaluación como una herramienta de gestión pública.

Israel es un caso donde se institucionalizó la evaluación de políticas durante la última década en una forma discontinua produciendo un sistema fragmentado. No es extraño que este sistema de evaluación esté dominado por motivos correspondientes a la primera etapa, es decir, de los administradores de programas que quieren mejorar sus políticas. En este sentido Israel recuerda el caso suizo que arrancó a finales de los años 80.

- En la mayor parte de los países implicados la evaluación institucionalizada, durante 1990, persistió y mantuvo su *status*; sin embargo, en los Estados Unidos, caso paradigmático de la primera etapa, la evaluación decreció; esto indica que el *status* que se ha alcanzado puede cambiar a lo largo del tiempo. A la vista del caso de los EE.UU. se puede indicar que los países europeos continentales tienden a mantener las institucionalizaciones incluso cuando están vaciadas de contenido. Los tres países nórdicos (Suecia, Noruega, Dinamarca), los Países Bajos y Alemania no cambiaron formalmente los dispositivos de evaluación en el gobierno nacional. Sin embargo, como se verá, hay indicaciones de cambios cualitativos también en estos países.

3. Influencia de la Unión Europea

El imperativo de evaluación establecido por la Unión Europea con respecto a los fondos estructurales estaba obligado a

tener implicaciones para los dispositivos nacionales de evaluación, en particular en los países que reciben enormes cantidades de los fondos europeos. El caso más destacado es Irlanda que estableció una unidad de evaluación controlada por el gobierno que coopera con direcciones generales en Bruselas. España es otro caso en el que la evaluación de políticas recibió un estímulo importante debido a la influencia de la Unión Europea; incluso en Alemania, desde 1990, el dinero de los fondos de cohesión que llegaba a Alemania Oriental ha incrementado el nivel de las actividades de evaluación en las instituciones ya existentes. Esto es así porque la Unión Europea funciona en un sistema de dos vías:

- Por una parte, la Corte Europea de Auditores (*European Court of Auditors*) coopera con los tribunales nacionales de cuentas (y en el caso de Alemania los auditores subnacionales) llevando a cabo comprobaciones legales y de cumplimiento;
- por otra parte, la inspección internacional de finanzas en la Comisión estimula una variedad de evaluaciones delegadas en varios países.

Con frecuencia se puede observar una tendencia a la burocratización de esta evaluación inducida externamente, en particular si se realiza en los tribunales tradicionales de cuentas (España y por otras razones: Corea).

4. Principales áreas de política para la evaluación

Se podría mantener que los sectores sanitarios y educativos constituyen las áreas dominantes donde se llevan a cabo evaluaciones de importancia, al menos en cierta medida. Además, los asuntos sociales, incluyendo la oferta de servicio social, es una de las áreas preferidas para llevar a cabo las evaluaciones (véase en particular el caso reciente de Israel). Estas tres áreas dominan con frecuencia los congresos nacionales de evaluación (por ejemplo, el Congreso Canadiense de Evaluación de 1997).

Además, los programas o las políticas orientados al cliente que están destinados a las minorías constituyen la realización de estudios de evaluación, por ejemplo, en Asuntos sobre Aborígenes en Canadá.

Además de estas «áreas de políticas blandas», los programas de alto coste, debido a sus particulares necesidades de legitimación, son un objetivo primordial de la evaluación. Mientras que la mayor parte de los gobiernos nacionales están recortando programas de gastos, la Unión Europea debe mencionarse en esta área como usuario e iniciador de las evaluaciones. Sin embargo, la estrecha relación entre los programas y el presupuesto va unida a una tendencia a poner en marcha instituciones auditoras tradicionales y transferirles la tarea de la evaluación. En

general, los auditores, sin embargo, no están preparados para llevar a cabo evaluaciones de calidad sino que mantienen su criterio auditor tradicional; esto puede mantenerse así porque la mayoría de las oficinas auditoras más importantes, que durante la primera fase se dedicaron a la evaluación de impacto (EE.UU. y Suecia) han recortado sus actividades de evaluación o han cambiado la calidad de las mismas hacia medidas de *output*.

5. ¿Existen cambios en la constelación de los actores?

Durante 1990 se habían creado dispositivos nacionales de evaluación en diversos países que se caracterizaban por comunidades de evaluación o actores dominantes que formaban la red de la evaluación. En general, estas constelaciones junto con el *status* alcanzado de la institucionalización no han variado, ni tampoco, los actores dominantes han desaparecido por completo. Sin embargo, han aparecido evaluadores en los niveles regionales y locales del gobierno, en particular en países de características federales o en los que existe una fuerte tendencia a la descentralización, como en España y Francia. De manera más pronunciada que hasta entonces, la evaluación de las políticas siguió siendo una función de la rama ejecutiva del gobierno. Esto se muestra de forma especialmente clara en el recién llegado, Israel, donde el Parlamento no está interesado en absoluto en la evaluación (posiblemente porque la función o se atribuye al auditor general neutral o se lleva a cabo en los departamentos de la rama ejecutiva relacionados con los programas). En Estados Unidos y en Canadá los legisladores parecen haber perdido interés en la evaluación de políticas.

Los auditores nacionales están jugando un papel ambivalente hoy día:

- En muchos países los evaluadores al máximo nivel se limitan a forzar o apelar a la rama ejecutiva para que incrementen el número de evaluaciones.
- En relación con esto la *General Accounting Office* de los EE.UU., aunque está de manera más clara ligada al ejecutivo, ya no parece que constituya una excepción y un contraejemplo con respecto a los auditores europeos.
- Es difícil decidir hasta qué punto los auditores al máximo nivel en Europa, que tuvieron su origen en el comienzo del Estado moderno en tiempos predemocráticos, están integrando la función auditora con la función evaluadora. Al menos en Alemania existe un sistema claro de dos vías que no implica a los auditores, incluso cuando están involucrados en la evaluación de políticas de la Unión Europea. El auditor general sueco comenzó, recientemente, a dedicarse a las medidas de *output*. En Israel, el auditor general tiene una tarea muy amplia, pero no realiza evaluaciones por sí mismo.

Como se ha mencionado anteriormente, las regiones en España y en Francia (y los *länder* orientales en Alemania) se han convertido en miembros de la comunidad de evaluación nacional porque son beneficiarios de los fondos estructurales de la Unión Europea.

Incluso estando institucionalizadas, la *cantidad y la calidad de las evaluaciones que se llevan a cabo pueden disminuir*. En los Estados Unidos, el declive de la evaluación en la rama ejecutiva, que ya empezó durante la presidencia de Reagan, continuó bajo el Presidente Clinton; y lo que es más sorprendente, la fuerza impulsora en el sistema de evaluación estadounidense, el Congreso junto con la *General Accounting Office* y su departamento de evaluación perdieron interés también en la función y han abandonado prácticamente las actividades evaluadoras del Congreso. La versión americana de la *New Public Management*, «reinventando el gobierno», puede haber contribuido a la pérdida de interés en la evaluación al poner énfasis en el *output* en lugar de en el impacto.

Desde Suecia se informa de una tendencia similar en donde la evaluación en la oficina central auditora se estaba desplazando también desde el impacto a la medida de *output*.

En este momento no se puede decidir si esto es una tendencia persistente o si se trata de que las instituciones de evaluación están, simplemente, capeando el temporal.

Al medir la capacidad nacional de evaluación, uno debe *tener cuidado de no propasarse*. En los lugares en donde está cambiando el equilibrio federal de tareas, como en los EE.UU., hay indicios de que los evaluadores que, anteriormente, estaban empleados por los gobiernos nacionales se están desplazando hacia los niveles subnacionales, hacia los gobiernos locales (Dinamarca) o están encontrando empleo en empresas privadas de consultoría (EE.UU y Suecia).

6. Cambios cualitativos en la evaluación de políticas

6.1. Consultas de gestión

En cierto número de países (Alemania, EE.UU., Suecia) hay fuertes indicios de que los consultores de la gestión están participando en el negocio de la evaluación. Esto no significa necesariamente que la evaluación no siga siendo llevada a cabo por instituciones especializadas del gobierno, aunque sea a un nivel inferior al que se hacía anteriormente.

Por razones obvias, la *emigración de la evaluación a niveles más bajos* del gobierno y/o la implicación de los consultores de gestión en el negocio de la evaluación cambia la naturaleza de la evaluación. Se observan dos tendencias:

- Se reemplazan las medidas de impacto por las medidas de *output*. Esto puede ser una consecuencia cuando se evalúan las políticas nacionales, en caso extremo, en el nivel del gobierno local porque aquí es difícil que la totalidad del programa pueda ser evaluada sino solamente los *outputs* locales de determinado programa nacional.
- Puesto que la implicación de los consultores de la gestión está ligada a la generalización del *New Public Management*, la concentración en los *outputs* también se debe a una tendencia a *evaluar las instituciones en lugar de los programas*. Aunque esto puede ser apropiado cuando están a cargo de un programa instituciones especializadas funcionalmente, si se evalúan autoridades multifuncionales con objetivos generales, tendrán que ser tenidas en cuenta en términos de varios programas y sus *outputs*. En tercer lugar, puesto que la descentralización es una característica principal de la *New Public Management* existe una tendencia tanto en EE.UU. como en el Reino Unido y en Alemania (en los sectores educativo y de investigación) para *reorganizar las instituciones relevantes* de tal manera que se conviertan en *autoevaluadoras*. De nuevo, es probable que estas instituciones reorganizadas se centren en los *outputs* en sus informes ordinarios.

No es de extrañar que cuando los evaluadores están preocupados con el *output* aparezca una búsqueda de medidas de calidad. Puesto que las medidas de *output* dependen fuertemente de datos cuantitativos, el problema de la calidad del *output* sigue siendo una preocupación central. Anteriormente, este aspecto se incluía en las evaluaciones de impacto de las políticas nacionales que se hacían a gran escala.

6.2. Popularización de la evaluación

Con el cambio desde la evaluación de programas a la evaluación de la organización (y ocasionalmente evaluación del personal) la noción clásica de la evaluación como centrada en el impacto queda supervalorada. En los países latinos el término evaluación es parte del lenguaje coloquial y anteriormente favorecía la tendencia a incluir en él no simplemente las evaluaciones *ex post*, sino también evaluaciones del tipo *ex ante* (análisis coste-beneficio). Hoy día, además de esta tendencia, el uso inflacionario de la palabra y el hecho de que haya entrado también en el lenguaje popular de otros países, la ha convertido en un término muy inclusivo, como la palabra «frustración» que ha proliferado desde los años 60 en el lenguaje coloquial en todo el mundo.

Con la descentralización organizativa en el sector público se puede decir que la presión para realizar auditorías externas ha aumentado. Un caso pertinente de interés puede ser Australia

donde en 1997 la *New Public Management* condujo a una estricta separación de la política y las operaciones. Como era de esperar, en las agencias operativas la medida del *output* aumentará mientras que en los ministerios encargados de la política, dada su racionalidad, deberán crear una demanda mayor de evaluaciones de impacto. Hasta el momento, sin embargo, nada ha cambiado en Australia. Además, se observa con frecuencia que con la privatización aumenta la regulación (por ejemplo, en el sector de las telecomunicaciones). Este nexo podría reforzar la tendencia a crear instituciones que se autoevalúen y así una nueva cultura de la certificación.

Desde el punto de vista de *los sumos sacerdotes de la evaluación* y de sus libros de texto originales, esto también se percibe

como una secularización del oficio. Para permanecer dentro del marco se podría argumentar que

- la responsabilidad tradicional de los ministros ante el Parlamento era una clase de confesión religiosa,
- con la evaluación de impacto el sacerdote fue sustituido por el psiquiatra,
- la irrupción de los consultores de gestión y los auditores privados en la emisión de certificados recuerda a la reciente tendencia popular en las sociedades avanzadas de «exhibirse» (preferiblemente en los *shows* televisivos).

Sin embargo, las personas siguen confesándose y los psiquiatras todavía no son superfluos.

Notas

Artículo traducido por Elisa ROMERO SEBASTIÁN.

* La argumentación está basada en Hans-Ulrich DERLIEN (1990), «Genesis and Structure of Evaluation Effort in Comparative Perspective» en Ray RIST (ed.), *Program Evaluation and the Management of Government: Patterns and Prospects Across Eight*

Nations, New Brunswick/London (Transaction Publishers), pp. 147-176, así como en un reciente documento de trabajo del grupo INTEVAL en Dublín (Mayo 1998) para poner al día el panorama de 1990.

** Profesor de la Universidad de Bamberg (Alemania).