

Estudio comparado sobre institucionalización de la evaluación en Europa y en América Latina

Blanca Lázaro

Colección **Estudios nº 15**

Serie **Estados de la cuestión**
Área **Finanzas Públicas**



Estudio comparado sobre institucionalización de la evaluación en Europa y en América Latina

Blanca Lázaro

Estudio nº 15

Serie: Estados de la cuestión

Área: Finanzas Públicas



PROGRAMA FINANCIADO
POR LA UNIÓN EUROPEA

Edita:

Programa EUROsociAL
C/ Beatriz de Bobadilla, 18
28040 Madrid (España)
Tel.: +34 91 591 46 00
www.eurosociasocial-ii.eu

Con la colaboración:

Fundación Internacional y para Iberoamérica
de Administración y Políticas Públicas (FIIAPP)



FIIAPP

COOPERACIÓN ESPAÑOLA



La presente publicación ha sido elaborada con la asistencia de la Unión Europea. El contenido de la misma es responsabilidad exclusiva de los autores y en ningún caso se debe considerar que refleja la opinión de la Unión Europea.

Edición no venal.

Realización gráfica:

Cyan, Proyectos Editoriales, S.A.

Madrid, julio 2015



No se permite un uso comercial de la obra original ni de las posibles obras derivadas, la distribución de las cuales se debe hacer con una licencia igual a la que regula la obra original.

Índice

Prólogo	5
I. Introducción	9
II. Marco conceptual de las estrategias para el impulso de la evaluación	11
1. ¿Qué es la evaluación de políticas públicas?	11
2. ¿En qué consiste la institucionalización de la evaluación?	15
III. La institucionalización de la evaluación en Europa.	27
1. La evaluación en los países nórdicos: el caso de Suecia.	27
2. La evaluación en los Países Bajos	33
3. La evaluación en el Reino Unido	41
4. La evaluación en Alemania	47
5. La evaluación en Francia	51
6. La evaluación en Bélgica	57
7. La evaluación en países de Europa meridional: Italia y España	61
IV. La institucionalización de la evaluación en América Latina	75
1. La evaluación en Chile	77
2. La evaluación en México	85
V. Conclusiones y recomendaciones	107
Anexo I. Bibliografía	117
Anexo 2. Referencias consultadas por país y tipo de fuente.	121

Prólogo

Una contribución a la literatura de las ciencias sociales sobre los sistemas de evaluación, tal y como la que se realiza en este estudio, representa para la FIIAPP un gran paso desde nuestro mandato fundacional de cooperar en el fortalecimiento de los sistemas públicos, mejorando sus estructuras, modernizando sus marcos normativos y sus políticas y capacitando a sus funcionarios y gestores. Tras unos inicios centrados en aportar capacidad de gestión en programas de hermanamientos de la Unión Europea, la FIIAPP ha ampliado paulatinamente sus proyectos de asistencia pública mediante la movilización de expertos y funcionarios de la Administraciones públicas en la ejecución de proyectos y programas en más de 90 países.

La FIIAPP ha tenido en este tiempo el privilegio de observar y participar activamente en la generación de valor público en varias regiones del mundo, y esto nos ha permitido comparar y extraer lecciones de experiencias muy diversas. Hoy por hoy, los sistemas de evaluación tanto en América Latina como en Europa se encuentran inmersos en procesos de cambio que están vinculados a la propia regeneración de los sistemas políticos, en todos los casos tratando de hacer más transparente la gestión pública y de rendir cuenta a una ciudadanía que utiliza nuevos canales de participación social. En algunos casos, son agencias especializadas quienes han ocupado el papel central desde el que se articulan las Agendas Nacionales de Evaluación, mientras que en otros la evaluación está permeabilizada en distintos Ministerios u Oficinas encargadas del presupuesto, la planificación o el gasto, siendo también muy diverso el papel que cumplen las instituciones como los Parlamentos en lo que se refiere al control y el seguimiento de las políticas públicas. En todos los casos, los sistemas se enfrentan al desafío de integrarse en nuevos modelos de gestión que mejoren los procesos de toma de decisiones, incorporando elementos que permitan desarrollar políticas, programas y proyectos poniendo el foco de atención exclusivamente sobre aquellos factores que las hacen pertinentes, eficaces y sostenibles.

Precisamente el programa de cooperación regional de la Unión Europea para la promoción de la cohesión social en América Latina, EUROsociAL II, coordinado por la FIIAPP, se ha marcado como uno de sus ejes principales de trabajo la mejora de la

calidad del gasto público mediante una mayor vinculación de los Planes y los Presupuestos y sobre todo de la integración de la evaluación en el ciclo de las políticas públicas, que son algunos de los principales retos que están enfrentando las administraciones públicas latinoamericanas para poder orientar su gestión hacia la obtención de resultados que satisfagan las necesidades de la ciudadanía.

Mediante las distintas actividades realizadas se ha perseguido un doble objetivo: la articulación y coordinación de los planes nacionales, sectoriales y regionales con el presupuesto público; y la incorporación de los resultados de la evaluación a los procesos de planificación, presupuesto y seguimiento.

En pos de este último objetivo, en julio de 2014 se celebró, en Montevideo (Uruguay), un Seminario Internacional de Evaluación, donde se intercambiaron experiencias entre los países participantes, con el objetivo de construir entre todos unas líneas básicas comunes de referencia para la institucionalización de la evaluación de las políticas públicas. Producto de este Seminario, en el que participaron representantes de Costa Rica, Ecuador, Paraguay, Perú y Uruguay, se elaboró el “Manual de Buenas Prácticas en la Sistematización de la Evaluación de Políticas Públicas. Experiencias de América Latina”.

Dando continuidad a este proceso, en marzo de 2015 se celebró en Lima, Perú, un taller en el que participaron instituciones clave en la gestión de la evaluación de políticas públicas de estos cinco países de América Latina, y donde se sentaron las bases para compartir hojas de ruta y crear un Grupo de Trabajo que permitiera a las instituciones fortalecerse recíprocamente, desde ámbitos concretos y en general con competencias homólogas. El estudio que se presenta a continuación significa una contribución a la formalización de este Grupo de Trabajo Interinstitucional de Evaluación de Políticas Públicas, y un acompañamiento al proceso de institucionalización de sus sistemas de evaluación, tanto en Europa desde la experiencia latinoamericana como en América Latina desde la experiencia europea. Compartimos una visión que pretende formalizarse en modelos de trabajo en red, produciendo estudios y metodologías sectoriales de evaluación sobre áreas estratégicas tales como el medioambiente, las infraestructuras o la seguridad ciudadana, aportando la experiencia de la administración pública, e introduciendo nuevas herramientas tales como la revisión entre pares (*peer review*) de evaluaciones que permitan generar nuevas miradas a los cambios que se están produciendo en una acción pública cada vez más interconectada regionalmente.

La FIIAPP tiene el compromiso de seguir apoyando estas tareas, y en este año 2015, declarado como *Año Internacional de la Evaluación* por la III Conferencia Internacional sobre Evaluación de Capacidades Nacionales de Sao Paulo, se han puesto en marcha una serie de actuaciones dirigidas a impulsar la gestión de la evaluación como uno de sus ejes centrales de trabajo e introducir la evaluación en los Programas

que gestiona como un aspecto necesario para generar evidencias y mejorar los resultados. La FIIAPP sigue a disposición de los organismos internacionales, y de aquellas instituciones nacionales que tengan un compromiso firme con la evaluación, en la consolidación de estos procesos a los que hoy contribuye EUROsociAL con este Estudio comparado sobre institucionalización de la evaluación en Europa y en América Latina.

Pedro Flores Urbano

Director de la FIIAPP

I. Introducción

El presente informe pretende contribuir a la descripción y análisis de los sistemas de evaluación de políticas públicas que están configurándose en diversos países europeos y de América Latina, así como a la identificación de factores determinantes de una correcta implantación de esos sistemas. Más allá del interés que esa descripción y análisis puedan tener en sí mismos, su objetivo último es servir de referencia para el planteamiento de posibles alternativas de institucionalización de la evaluación en cuatro países de América Latina: Costa Rica, Ecuador, Perú y Paraguay.

La emergencia, configuración y evolución de sistemas de evaluación en ámbitos estatales es objeto de análisis desde hace relativamente poco tiempo, lo que constituye sin duda una limitación para este estudio. Ello afecta tanto a qué debe entenderse por sistema de evaluación como a cuáles son sus componentes, cuál es su lugar y su misión en un entorno institucional público, qué recursos, qué actividades, qué resultados deben esperarse de dicho sistema; cuáles son los factores que determinan su buen o mal funcionamiento, la obtención de resultados satisfactorios o deficientes.

Por otra parte, es importante observar el entorno social y cultural en que la evaluación se desarrolla. Recetas exitosas en un contexto determinado pueden no funcionar en otro distinto ya que, a fin de cuentas, deberán ser implementadas por otras personas, inmersas en una cultura política, administrativa y organizativa diferente; en unas condiciones sociales y económicas que pueden variar substancialmente, y formando parte de unas redes personales, profesionales y configuraciones institucionales también distintas.

El informe se ha realizado a partir de la revisión de literatura académica y de documentos producidos por organismos internacionales y por organismos públicos de los países estudiados. No ha incluido trabajo de campo, y en particular entrevistas con responsables u otros actores implicados en los sistemas de evaluación estudiados, por lo que no ha sido posible contrastar la información documental, ni actualizarla en algunos casos. En el caso de España, al análisis documental se añade la experiencia de quien redacta el informe en la dirección de un organismo de referencia en evaluación de políticas en ese país, entre 2008 y 2013.

Aún con esas limitaciones, el informe aporta el valor de reunir en un único documento una visión panorámica de la evolución y de la situación actual de los sistemas de evaluación en buena parte de Europa y en una selección de países de América Latina. Y de analizar los principales resortes, obstáculos y lecciones aprendidas respecto a su surgimiento y desarrollo.

El texto se organiza en cuatro apartados. El primero revisa y expone el marco conceptual actual sobre los sistemas de evaluación y las estrategias para impulsarlos. El segundo se centra en los sistemas de evaluación en países europeos. El tercero analiza la situación en algunos países de América Latina. Finalmente, se formulan conclusiones y recomendaciones con el ánimo de que sirvan de base para trazar a corto plazo hojas de ruta específicas para la institucionalización de la evaluación en los cuatro países mencionados.

II. Marco conceptual de las estrategias para el impulso de la evaluación

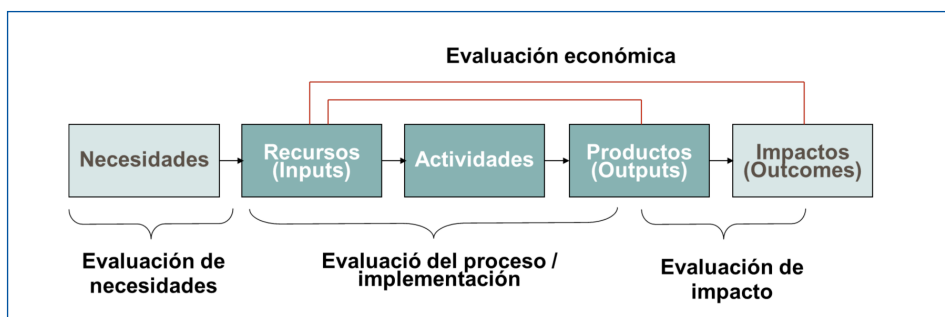
1. ¿Qué es la evaluación de políticas públicas?

Entre las numerosas definiciones existentes en la literatura sobre la evaluación de políticas públicas, en este estudio tomaremos como referencia la incluida en la comunicación de la Comisión Europea *Responding to strategic needs: reinforcing the use of evaluation*, según la cual:

“...evaluation is a judgment of interventions according to their results, impacts and needs they aim to satisfy. It is a systematic tool which provides a rigorous evidence-base to inform decision-making.” (Commission of the European Communities, 2007, pág. 3)

Esa definición nos parece especialmente adecuada porque incluye diferentes perspectivas de evaluación de políticas públicas (Figura 1) a la vez que pone énfasis en necesidades e impactos. Es decir, promueve como criterio principal de evaluación conocer hasta qué punto la intervención pública da respuesta a necesidades sociales.

Figura 1. Enfoques de evaluación de políticas públicas.



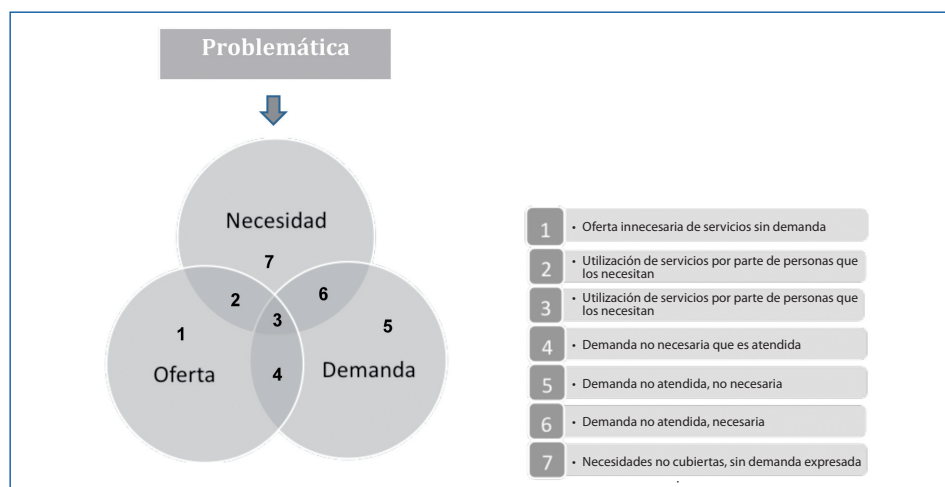
Fuente: Instituto Catalán de Evaluación de Políticas Públicas (Ivàlua).

Ello implica, por un lado, partir de la identificación de las necesidades o problemáticas (sociales, económicas...) que originan la intervención pública. Algo nada banal, dado el reto

conceptual y analítico que comporta definir y estudiar problemas relacionados con la “calidad de vida”, la “cohesión social”, la “competitividad de la economía”, la “desafección ciudadana”, etc., tantas veces presentados en el discurso político como desafíos a abordar. Se suma también a esa dificultad la confusión entre necesidades y demanda de servicios públicos, una confusión frecuente en enfoques gerenciales de reforma o de modernización administrativa, atentos a las expectativas y demandas de los “clientes” o “usuarios” de los servicios públicos, no necesariamente coincidentes con sus genuinos destinatarios (Figura 2).

Por otro lado, según esa definición, la evaluación de políticas públicas implica esclarecer la vinculación o relación causal entre la intervención pública y los cambios que persigue respecto a esa situación de necesidad, tanto en un plano teórico –objeto de la evaluación del diseño de la intervención– como empírico –objeto de la evaluación de impacto o efectividad–. Ese análisis de causalidad diferencia la evaluación de actividades que a menudo se confunden con ella, como el monitoreo o seguimiento de políticas públicas, la auditoría o la valoración del rendimiento.

Figura 2. Necesidad, oferta y demanda de servicios.



Fuente: Casado, 2009 (adaptación de William y Wright, 1998).

Veamos brevemente las características de ese otro tipo de actuaciones, para distinguirlas más claramente.

- La monitorización del rendimiento o *Performance Monitoring* se desarrolla especialmente a partir de la ola de reformas administrativas y de gestión pública acaecida desde finales de los años 80, bajo el paradigma del *New Public Management* (NPM) (Osborne y Gaebler, 1992). Vinculados a la gestión por resultados (*Results-based Management*), esos sistemas se conciben como instrumentos de apoyo a la dirección política y a la dirección y gestión de organizaciones públicas, se desarrollan normalmente

desde el interior de las propias administraciones públicas, aunque incluyen a menudo mecanismos de comparación o *Benchmarking* con otras organizaciones similares, rankings contruidos en función de determinados indicadores de rendimiento o de calidad de los servicios, etc.

Una diferencia importante entre evaluación de políticas públicas y medición del rendimiento, más allá de cuestiones metodológicas, es que la primera se centra normalmente en “políticas substantivas”, definidas como

“...the ‘normal’ policies (such as social policy, employment policy, youth policy, housing policy) which essentially target the socio-economic situation and development in the policy environment” (Wollman, 2003, pág. 18).

Mientras que la medición del rendimiento se centra más bien en aspectos de implementación y resultados, a menudo a nivel de organización pública y no necesariamente de política pública. Evaluación y medición del rendimiento no son actividades excluyentes, sino complementarias, con sinergias evidentes aunque poco explotadas, que normalmente se implantan y progresan, como veremos, por vías paralelas. La Tabla 1 muestra de forma sintética las diferencias entre ambas aproximaciones.

Tabla 1. Diferencias entre Evaluación y Performance Monitoring.

	Evaluación	Performance Monitoring
<i>Objeto</i>	Políticas públicas / programas	Organizaciones
<i>Tiempo</i>	Episódica, estudios puntuales	Práctica regular, periódica, con seguimiento como mínimo anual
<i>Métodos</i>	Combinación de métodos cuantitativos y cualitativos	Mayoritariamente cuantitativos
<i>Quién lo impulsa</i>	Órganos o unidades responsables de políticas públicas	Órganos o unidades a cargo del control presupuestario, organización y/o RRHH
<i>Quién lo lleva a cabo</i>	Expertos externos y/o unidades internas de evaluación	Órganos o unidades administrativas.

Fuente: elaboración propia a partir de Talbot (2010).

- El monitoreo o seguimiento de políticas públicas: las diferencias entre esta técnica y la evaluación son similares a las comentadas en el caso del *Performance monitoring*, con el matiz añadido de que ahora la función de monitoreo y seguimiento se centra en las políticas y programas, en lugar de en las organizaciones. Así, en este caso, el monitoreo o seguimiento proporciona información sobre la situación de la intervención y de su contexto. Permite caracterizar el problema de partida y cómo evoluciona, así como el funcionamiento de la intervención pública, su cobertura, sus resultados –siempre y cuando los indicadores estén cuidadosamente seleccionados, definidos y medidos-. Sin embargo, el monitoreo no contempla el análisis en profundidad de esos diferentes componentes, ni permite

determinar causalidad entre intervención y problema, tal como pretende hacer la evaluación de efectividad.

- Las prácticas de *Performance Audit, Inspection and Oversight* (Auditoría del rendimiento, inspección y supervisión). Surgidas también al calor del movimiento de NPM, las actividades en este caso se orientan a verificar y controlar si las organizaciones públicas y sus responsables han realizado las misiones que tienen encomendadas, y si lo han hecho mediante un uso eficiente de los recursos. Por tanto, se trata de un tipo de análisis no tanto vinculado a verificar la efectividad o efecto de las actividades o programas, sino a los logros alcanzados por las organizaciones y sus responsables y al uso eficiente de los recursos que han gestionado. Este tipo de análisis se refuerza en la etapa del NPM, como contrapartida necesaria a una mayor autonomía de los gestores públicos, que va de la mano con una mayor necesidad de rendición de cuentas y de control. El ejercicio de estas prácticas normalmente reside en tribunales de cuentas, *National Audit Offices* u organismos equivalentes, lo que explica el auge de este tipo de instituciones en países de la OCDE –y en particular el de sus actividades de *performance auditing*–, especialmente a partir de la década de los 90.
- El análisis de impacto regulatorio o *Regulatory Impact Analysis*: se trata de un análisis ex ante de las alternativas y de los posibles efectos de disposiciones normativas, en términos de efectividad y de coste efectividad (OECD, 2009). Presta especial atención a los costes de todo tipo que las normas puedan suponer para el sector privado, y en particular para el sector privado empresarial. Para ello se emplean técnicas de análisis de las cargas administrativas, como el *Standard Cost Model*.
- La acreditación de servicios u organizaciones. Muy presente también en países de la OECD, en ámbitos de servicios públicos, como educación superior, prisiones, hospitales, etc. Se lleva a cabo por parte de organismos públicos (p.ej., agencias de evaluación de las universidades o de la investigación) o privados (empresas de consultoría que implantan y/o acreditan sistemas de gestión de la calidad en servicios públicos diversos), en el marco de sistemas de gestión de la calidad.
- Finalmente, cabe mencionar prácticas mixtas de monitoreo y evaluación (M&E). Esas prácticas tratan de superar las limitaciones del seguimiento estricto de programas –falta de profundidad en el análisis, excesivo foco en procesos y no en resultados e impactos– y de la práctica evaluativa aislada –estudios puntuales, no siempre adaptados al calendario y forma de presentación de resultados que maximice su utilización. La idea es obtener un flujo constante de información y conocimiento riguroso, bajo formas diversas –indicadores, evaluaciones ex ante, de implementación, de impacto– adaptadas a cada fase del ciclo de las intervenciones públicas y a sus procesos de gestión y de toma de decisiones, incluyendo la asignación de recursos presupuestarios (Rist y Stame, 2011). Algunos países de América Latina, como Chile, constituyen un referente internacional en el desarrollo de ese tipo de sistemas.

2. ¿En qué consiste la institucionalización de la evaluación?

Entendemos la institucionalización de la evaluación como un proceso tendente a establecer un sistema de evaluación en entornos gubernamentales, a través de políticas o estratégicas concretas. En este apartado veremos primero qué debe entenderse por sistema de evaluación, y a continuación trataremos de caracterizar conceptualmente las políticas o estrategias necesarias para impulsarlo.

Es necesario, sin embargo, señalar primero dos limitaciones de la evaluación, una técnica y otra política, a tener muy en cuenta en cualquier estrategia de desarrollo e implantación de la misma. De otro modo, pueden generarse falsas expectativas sobre lo que la evaluación puede ofrecer y sobre el rol que debe jugar en un sistema político democrático, que resulten contraproducentes para una implantación exitosa.

Las limitaciones de la evaluación

Partimos del supuesto de que el interés en impulsar la evaluación se basa en su capacidad de generar evidencias robustas sobre el funcionamiento, efectividad y eficiencia de las políticas públicas que retroalimenten los procesos de toma de decisiones y de esa forma consigan incrementar la efectividad de las intervenciones públicas.

Sin embargo, desde un punto de vista científico-técnico, esa producción de evidencias es un proceso continuo donde los informes de evaluación raramente aportan evidencias del todo concluyentes, y por tanto no deberían considerarse un punto final. Bien al contrario, cada informe de evaluación suele responder una serie de preguntas pero plantea al mismo tiempo otras nuevas; sugiere diseños o variantes del programa que, de llevarse a la práctica, será necesario evaluar; propone mejoras en los sistemas de información que permitirán llegar más lejos en futuros análisis. En cambio, demasiado a menudo se espera que los informes de evaluación sean dictámenes rotundos, definitivos, sobre la idoneidad de un programa. Esa expectativa puede generar frustración, que a su vez puede redundar en una percepción negativa de la evaluación y dificultar su avance.

Por otro lado, desde un punto de vista político, las evidencias que produce la evaluación deben sin duda estar presentes en las mesas donde se toman las decisiones sobre la continuidad, modificación o supresión de los programas. No obstante, en palabras de Rist (2006) en un sistema democrático no deben aspirar a ocupar el lugar de honor –ya que ello sería más propio de sistemas tecnocrático-autoritarios–, sino a ser tenidas en cuenta junto con criterios legítimos de carácter político.

En definitiva, si se asumen esas características y esas limitaciones de la evaluación, la estrategia de desarrollo e implantación que se adopte debe acompañarse necesariamente de un proceso de cambio hacia una Administración diferente, más reflexiva y más consciente de la complejidad de la realidad social y de las propias limitaciones

para hacerle frente. Y más proclive, por tanto, a apoyar la generación continua de conocimiento de calidad, a promover el debate político y social bien informado sobre las posibles alternativas de intervención y sus resultados, y a afrontar el fracaso y la crítica como una forma de aprendizaje. Tal como dice Innerarity:

"La democracia es un sistema político que genera decepción... especialmente cuando se hace bien. Cuando la democracia funciona bien se convierte en un régimen de desocultación, en el que se vigila, descubre, critica, desconfía, protesta e impugna." (Innerarity, 2015).

Los sistemas de evaluación

La necesidad de desarrollar sistemas de evaluación con vistas a la regularización de su práctica y de su utilización en entornos gubernamentales empieza a ser objeto de estudio a finales de la década de los 90, a través de trabajos como los de Boyle, Lemaire y Rist (1999). Estos autores señalan que un "régimen de evaluación" es aquel interesado en la configuración de la capacidad para la evaluación, en la práctica de la evaluación y en aspectos organizativos e institucionales relacionados.

Por otra parte, según Furubo y Sandahl, la existencia de un sistema de evaluación se da.

"...when evaluations are no longer commissioned and conducted on an 'ad hoc' basis but through more permanent arrangements, which aim to guarantee, in advance, the supply of evaluative information... [and] at the same time, the evaluations conducted are put to suitable use." (Furubo y Sandahl, citados por Leeuw y Furubo, 2008, pág. 159).

Varone y Jacob (2004) abundan en esa concepción de los sistemas de evaluación, así como Barbier y Hawkins:

"...practices over time tend to be transformed into routines, especially when they succeed, and they are crystallized in formal institutions." (Barbier y Hawkins, 2012, pos. 158¹)

En suma, podemos definir un sistema de evaluación de políticas públicas como aquel en que ésta forma parte habitual del ciclo de vida de las políticas y programas públicos, se lleva a término de forma sistemática y metodológicamente rigurosa, sus resultados son utilizados por parte de decisores políticos y de gestores a cargo de las intervenciones, y sus resultados son accesibles a la ciudadanía.

En tal sistema se entrecruzan, por tanto, valores, prácticas e instituciones propias de un determinado sistema político y administrativo (Figura 3).

1. Nota: al tratarse de un libro electrónico, sustituimos la referencia a la página por la referencia a la "posición" indicada en el dispositivo de lectura

Figura 3. Sistemas de evaluación.



Fuente: adaptación de Barbier y Hawkins (2012).

Furubo y Sandahl (2002) analizan los sistemas de evaluación en 21 países y diversos organismos internacionales, sobre la base de una serie de indicadores que posteriormente completan Leeuw y Furubo (2008) y que pueden agruparse en las siguientes categorías: existencia de una base epistemológica común; existencia de una cultura favorable a la evaluación; grado de institucionalización, organización y demanda de evaluación; oferta de evaluación; práctica evaluadora y grado de utilización y difusión de la evaluación, a efectos de mejora de los programas y de *accountability* o rendición de cuentas. Veamos brevemente esos rasgos distintivos:

- Presencia de la evaluación en el discurso político, administrativo y social. En particular, sobre la necesidad y la utilidad de evaluar, sobre cómo organizar la práctica evaluadora, cómo utilizar y difundir los resultados de las evaluaciones, etc.
- Existencia de una *base epistemológica común*: se refiere a la necesidad de consenso entre los diferentes actores presentes en el sistema de evaluación sobre qué es la evaluación y qué tipo de conocimiento debe producir (p.ej. ¿evaluaciones de efectividad o impacto? ¿con métodos experimentales o cuasi experimentales? ¿*Performance auditing*? ¿Monitoreo de programas y proyectos? ¿*Regulatory Impact Assessment*?, ¿cumplimiento de objetivos y/o de procedimientos? ¿Causalidad entre intervenciones y resultados?....; ¿Cómo debe llevarse a cabo? y ¿con qué finalidad?). Al respecto, Jacob (2005) distingue diferentes marcos cognitivos que encuadran los procesos de institucionalización de la evaluación, y que se encuentran normalmente mezclados, no en “estado puro”, en los diferentes países en los que se observa el proceso:
 - Una finalidad clásica de “control” de la Administración.
 - Una orientación “gerencialista”, basada en la medición del rendimiento y de los efectos (asociada, como ya se ha visto, a la corriente del *New Public Management*).
 - Una orientación “reflexiva”, donde la evaluación aparece como un modo alternativo de transmisión de la información sobre las políticas públicas y sobre la gestión

pública, dirigido a los decisores públicos, a los actores que participan en la implementación de la política, a los beneficiarios y a los ciudadanos en su conjunto. Este es un punto clave en el desarrollo de la evaluación: la clarificación de estos marcos, de lo que se persigue en cada caso, de modo que puedan establecerse “comunidades de práctica evaluadora” que hablen un lenguaje similar, que compartan una misma “cultura” de evaluación, tal como la define Speer:

“... evaluation culture is ... a mainstream preference for specific evaluation practices, approaches and systems. Evaluation cultures develop and are influenced by the public policy sector and professionals and they are based on collective and operational choices as well as by rules-in-use.” (Speer, 2012, pos. 1407)

- *Responsabilidad organizativa*: deben existir organizaciones (una o varias) que realicen evaluaciones, no sólo evaluadores individuales. Pero debe haber no solo organizaciones que produzcan o que tengan el encargo de producir evaluaciones, sino también organizaciones que las soliciten y que las utilicen, que coordinen la producción y la utilización, etc.

Al respecto, Jacob (*op. cit.*) distingue entre diversas configuraciones organizativas posibles:

- *Monopolio*, donde una única organización ostenta el rol dominante en el impulso y en la práctica de la evaluación. En la literatura se señala este sistema como propio de etapas incipientes de institucionalización de la evaluación.
- *Pluralismo centralizado*: en el seno de un dispositivo organizativo de evaluación, un organismo y/o u conjunto de normas ocupa una plaza preponderante, determinando el comportamiento del resto, que se someten voluntariamente a ese sistema. Esta configuración es bastante más común, especialmente en sistemas maduros de ámbito anglosajón².
- *Concurrencia fragmentada*, donde no existe el elemento centralizador del caso anterior. Puede darse en países fuertemente descentralizados o bien con desarrollos incipientes y débiles de la institucionalización de la evaluación³.
- *Permanencia*: las actividades de evaluación no deben ser puntuales, aisladas, sino que deben realizarse de forma regular, continuada, y deben alcanzar un cierto volumen y cobertura. Todo ello implica contar con recursos presupuestarios asignados también de forma explícita y ordinaria.
- *Énfasis en la utilización*: los resultados de la evaluación deben de estar vinculados con los procesos de toma de decisiones y procesos de implementación de los programas (asignación de recursos, ajustes en objetivos o actividades, etc.).

2. Algunos ejemplos destacados, como es el caso de los sistemas de evaluación federales en Estados Unidos y Canadá, quedan fuera del ámbito de este estudio.”

3. Desde nuestro punto de vista, esta clasificación debería incluir una configuración pluralista “descentralizada”, situada entre el pluralismo centralizado y la concurrencia fragmentada, para acoger sistemas de evaluación desarrollados pero sin un núcleo coordinador. Mientras que el escenario de concurrencia fragmentada podría reservarse a sistemas de evaluación todavía débiles que aspiran según los casos a una configuración pluralista o bien pluralista centralizada sin conseguirlo todavía plenamente.

- La *difusión pública* de los resultados de las evaluaciones, como medida de transparencia y de *accountability*. Tal como recuerda Chelimsky:

"The government's responsibility to the people for its actions naturally includes telling them what it's been doing. And accountability also requires evaluators to report their findings to the public." (Chelimsky, 2008, pág. 402)

Por otra parte, los sistemas de evaluación de políticas públicas no se articulan siempre en -ni coinciden necesariamente con- marcos institucionales clásicos propios de sistemas políticos:

- Pueden desarrollarse alrededor de proyectos específicos, ya sea transnacionales, como es el caso de la *Campbell Collaboration*⁴ centrado en evaluaciones experimentales o cuasi experimentales y que ha llegado a constituir un sistema propio de evaluación (Leeuw y Furubo, op. cit.). O la *International Initiative for Impact Evaluation* (3iE)⁵, en el ámbito de los programas de desarrollo.
- O bien, alrededor de sectores específicos de política pública, como en experiencias que veremos cuando repasemos la situación de la evaluación en diferentes países europeos, y que mencionamos aquí sucintamente: la *Education Endowment Foundation*, en el Reino Unido, el *Fonds d'Expérimentation pour la Jeunesse*, en Francia o, con una arquitectura institucional más difusa, propia, como veremos más adelante, del modelo alemán, el sistema de evaluación de políticas activas de empleo generado a partir de las reformas Hartz en 2003 (Jacobi y Kluve, 2006).
- O, en estrecha relación con las experiencias mencionadas en el punto anterior, con nuevas tendencias que vinculan innovación y evaluación de impacto, como los Social Impact Bonds, en Reino Unido⁶ o Estados Unidos⁷.
- O bien, dentro de una misma organización, pueden coexistir diferentes subsistemas de evaluación. Sería el caso, por ejemplo, de los impulsados por la Oficina de Evaluación y Supervisión (OVE) y por la Oficina de Planificación Estratégica y Efectividad en el Desarrollo (SPD) en el BID; o por el *Independent Evaluation Group* (IEG) y la *Development Impact Evaluation Initiative* (DIME) en el Banco Mundial), etc.
- O bien dentro de un mismo sistema político pueden coexistir diferentes sistemas o subsistemas de evaluación de ámbito estatal, regional o local.

Factores que influyen en el desarrollo y evolución de los sistemas de evaluación

Furubo y Sandahl (2002) identifican una serie de factores determinantes en el desarrollo de los sistemas de evaluación:

4. www.campbellcollaboration.org

5. <http://www.3ieimpact.org/>

6. <https://www.gov.uk/social-impact-bonds>

7. <http://www.whitehouse.gov/omb/factsheet/paying-for-success>

- En primer lugar, la *calidad democrática*, y en particular el arraigo de prácticas de transparencia y de participación que favorezcan el debate social abierto sobre las actuaciones públicas.
- En segundo lugar, la *tradición científica, técnica y de gestión pública*, que si es sólida tenderá a favorecer aproximaciones racionalistas a la toma de decisiones. Por el contrario, cuando la tradición jurídico-administrativa es dominante en la Administración y el conocimiento académico se encuentra atomizado y disperso, el progreso de la evaluación será más difícil (Bemelmans-Videc et al., 1994). Al respecto, Jacob (*op. cit.*) añade la importancia de favorecer la “hibridación” entre las culturas y las prácticas administrativas y académicas, así como entre áreas especializadas de conocimiento académico. Esa hibridación sería fundamental para el desarrollo de la evaluación en las administraciones públicas.
- La *elevada inversión pública en determinados sectores* propios del estado del bienestar –o, podríamos añadir en el contexto actual, la necesidad perentoria de ajustes– favorece asimismo el uso de la evaluación en la toma de decisiones.
- Finalmente, contribuye también la *existencia de fuerzas exógenas al sistema y favorables a la evaluación*. En particular, la influencia de organismos internacionales de cooperación al desarrollo y, en el ámbito de la Unión Europea, las obligaciones de evaluación asociadas al acceso a fondos estructurales.
- Por otra parte, algunos autores señalan también el *contexto institucional y características propias del sistema político*. Al respecto, según Varone y Jacob (*op. cit.*), la evaluación se institucionalizaría más fácilmente en democracias con sistemas electorales mayoritarios o regímenes presidencialistas que en democracias con sistemas proporcionales o de tipo consociativo, ya que en estas últimas el recurso a criterios estrictamente políticos para llegar a acuerdos puede relegar el papel de la evaluación en los procesos de toma de decisiones⁸. Bajo esta misma perspectiva institucional, pero referida ahora a los *órganos de control*, los mismos autores estiman que el grado de independencia de éstos respecto al poder ejecutivo puede ser un elemento facilitador de la institucionalización. Y si además desarrollan prácticas de evaluación, pueden contribuir también a activar la función de control del Parlamento, que se servirá de ellas con esa finalidad, o bien desarrollará sus propios dispositivos de evaluación.
- Los *procesos de reforma de la Administración*, aunque ya hemos visto que en algunos casos pueden promover, bajo el epígrafe de la evaluación, iniciativas diversas y que no siempre contribuyen al desarrollo de ésta. Así, Wollmann (*op. cit.*) considera que la emergencia de procesos de reforma del sector público es probablemente un factor explicativo importante de la aparición de prácticas de evaluación interna. En cambio, con respecto a la evaluación externa independiente, considera que la existencia de actores institucionales como el Parlamento

8. Sin embargo Suiza, un país que queda fuera de este informe y que es un modelo de democracia consociativa, cuenta con un sistema de evaluación robusto.

o el Tribunal de Cuentas, así como la madurez de la cultura política del país, tiene probablemente una mayor incidencia.

Para apreciar la influencia de algunos de esos factores en la emergencia de sistemas y prácticas de evaluación en el entorno internacional, a través de los Cuadros 1, 2, 3 y 4 ofrecemos una visión panorámica de la evolución de la evaluación.

Cuadro 1. Principales etapas en la expansión de la evaluación. Primera ola: la expansión del Estado del Bienestar período de entreguerras y años 50 y 60

Características del entorno social y político:

- Desarrollo del Estado del Bienestar en Europa y Estados Unidos. Creciente complejidad de la actuación gubernamental y de la dirección y gestión pública, debido a la puesta en marcha de grandes programas públicos de asistencia social, salud y educación.
- Entorno de gestión pública centralizada, con fuerte presencia de planificación estratégica y operativa asociada al despliegue del estado del bienestar y de los grandes programas de lucha contra la pobreza, educación, salud, planificación urbana....
- Necesidad de legitimidad política de la intervención pública en esos nuevos ámbitos.
- Desarrollo de las ciencias, y en particular de las ciencias sociales y del análisis de políticas públicas.
- Ideal de hacer el gobierno más racional, más científico y en última instancia más eficaz para hacer frente a problemáticas sociales.
- Separación entre toma de decisiones, conceptualización de la política e implementación. La evaluación afecta a la primera; la segunda se concibe de forma mecanicista, sin implicaciones a tener en cuenta para la evaluación y conceptualización de las políticas.

Características de la evaluación en esa etapa:

- Centrada en analizar resultados e impactos de políticas substantivas, relativas a educación, servicios sociales, salud, etc.
- Institucionalización (prácticamente) inexistente.
- Práctica a través de estudios puntuales encargados a expertos externos académicos.
- Utilización de métodos eminentemente científico-experimentales, cuantitativos.
- Objetivo de informar el proceso de decisiones públicas sobre los medios más idóneos para lograr fines fijados en el nivel político-administrativo.

Fuente: elaboración propia a partir de Furubo y Sandahl (2002) y Vedung (2010).

Cuadro 2. Principales etapas en la expansión de la evaluación. Segunda ola: freno a la expansión del estado del bienestar y avance del relativismo cultural y científico.

Características del entorno social y político:

- Fuerte crisis económica, motivada por el incremento del precio del petróleo desde principios de los 70 y a lo largo de toda la década.
- Las repercusiones de la crisis en las haciendas estatales empiezan a erosionar el Estado del Bienestar.
- Inestabilidad política internacional (guerra del Yom Kippur, revolución islámica en Irán, instauración de dictaduras en diversos países de América Latina) aunque con evoluciones también positivas, como la transición a la democracia en España.
- Inicio de procesos de desconcentración y de descentralización administrativa en algunos países.
- Erosión del paradigma científico positivista, a favor de enfoques constructivistas, que niegan la existencia de una verdad objetiva.

Características de la evaluación en esa etapa:

- Redefinición del papel de la evaluación, que empieza a orientarse a la reducción de las intervenciones públicas y a la maximización de la eficiencia en el uso de los recursos públicos.
- El mismo tiempo, emergen nuevas demandas de participación por parte de grupos sociales (jóvenes, mujeres, minorías étnicas) y usuarios de servicios públicos, y de más diálogo y comunicación en las evaluaciones (“Dialogue-oriented wave”, Vedung, 2010), en un entorno de reivindicación de mejoras en el funcionamiento de los sistemas democráticos representativos.

Fuente: elaboración propia a partir de Furubo y Sandahl (2002) y Vedung (2010).

Cuadro 3. Principales etapas en la expansión de la evaluación. Tercera ola: de los 80 a mediados de los 2000. Auge del neoliberalismo y New Public Management

Características del entorno social y político:

- Encadenamiento de períodos de crisis y de bonanza económica –globales (74-75, 92, 2007), regionales o sectoriales (96-97, crisis financiera de los “tigres” del sudeste Asiático; crisis de las “punto.com”, etc.), de país (reestructuraciones económicas en UK, España, etc., en los 80)- y de retroceso del Estado del Bienestar.
- Ampliaciones sucesivas de la Unión Europea (+15, +25, +27).

Cuadro 3. Principales etapas en la expansión de la evaluación. Tercera ola: de los 80 a mediados de los 2000. Auge del neoliberalismo y New Public Management (cont.)

- Incremento de programas de asistencia técnica con vistas a la incorporación de nuevos países a la Unión Europea, así como de fondos estructurales para promover el desarrollo y la equiparación entre países miembros.
- Oleada de reformas en las administraciones públicas conocidas como New Public Management, en un entorno de cuestionamiento del papel del estado en la sociedad y de los contornos de la Administración pública y de los servicios públicos. Las reformas afectan de forma desigual –más intensa en el ámbito anglosajón– a países de la OCDE y se extienden también a otras áreas. En Europa se inicia en los 80, en el Reino Unido y países escandinavos. Posteriormente, en los 90, comienza a extenderse por Europa continental, y en los 2000 alcanza los países de Europa meridional.
- Simplificando mucho, en el NPM, los mercados y la competencia son generalmente preferidos como forma de provisión de servicios públicos de forma más eficiente a unos beneficiarios concebidos como clientes. Se diluye la separación clásica entre dirección política y administración burocrática, otorgándose a los gestores públicos una gran autonomía que incluye la fijación de objetivos y la fijación de instrumentos administrativos y de gestión para llevarlos a cabo.
- Desconfianza en el papel de la planificación centralizada. Desregulación, privatización, eficiencia y orientación al cliente se convirtieron en nuevos términos fetiche.

Características de la evaluación:

- El monitoreo y la medición del rendimiento (Performance measurement) a través de baterías de indicadores centrados fundamentalmente en inputs, actividades y outputs de organizaciones a cargo de funciones administrativas y de servicios públicos, se van generalizando como instrumento de rendición de cuentas, ya sea desde entes territoriales a entes centrales, entre éstos y agencias ejecutivas, o entre proveedores privados de servicios públicos y las autoridades competentes.
- La evaluación de la efectividad de las políticas públicas queda en un segundo plano. En cambio, se enfatiza la valoración de la satisfacción de los clientes y usuarios de servicios públicos, el benchmarking, el análisis de la eficiencia organizativa e individual. En definitiva, la finalidad gerencial del NPM, orientada a mejorar la eficacia interna, tiende a disminuir el peso de la evaluación.
- Auge, al final de la etapa, y de la mano de políticas de Better Regulation, de la evaluación ex ante de impacto y de eficiencia (Impact Assessment, Regulatory Impact Assessment).

Cuadro 3. Principales etapas en la expansión de la evaluación. Tercera ola: de los 80 a mediados de los 2000. Auge del neoliberalismo y New Public Management (cont.)

- Por otra parte, se inician procesos de institucionalización: en la Unión Europea, asociados sobre todo a los Fondos estructurales y el Fondo Social Europeo, así como a procesos de reforma administrativa. En América Latina, vinculados en algunos casos también a esos mismos procesos, aunque también debido a factores exógenos, como la evaluación vinculada a grandes programas sociales financiados por organismos internacionales.
- En ambos casos, el impulso y gobierno de esos procesos de institucionalización proviene a menudo de ministerios o departamentos de Finanzas, responsables de la elaboración y la gestión del presupuesto, aunque también de Departamentos de la Presidencia y/o Ministerios sectoriales en algunos casos.

Fuente: elaboración propia a partir de Furubo y Sandahl (2002) y Vedung (2010).

Cuadro 4. Principales etapas en la expansión de la evaluación. Cuarta ola: finales 2000 – actualidad (2014)

Características del entorno social y político

- Gran crisis financiera y económica global.
- Crisis fiscal del Estado, especialmente en los países del Sur de Europa.
- Ola global de políticas de austeridad y recesión económica en diferentes regiones y países.
- Agudización de la desigualdad social a escala mundial.
- Crisis institucional y social, con cuestionamiento creciente del papel del Estado, de la clase política, desconfianza en el sector privado y en la connivencia entre intereses económicos y clase política.
- Surgimiento, ya desde principios de 2000 y acentuándose al final de la década, tendencias críticas con el NPM (New Public Governance, Public Values Governance, New Public Service...). Esas nuevas tendencias se plantean no sólo cómo gestionar mejor, sino, de forma más amplia, cómo gobernar en unos entornos sociales cada vez más diversos, donde hay que hacer frente a problemas de complejidad creciente que afectan no solo a los gobiernos, sino también a las empresas, al tercer sector y a la sociedad en general. Y que conviven con la desconfianza cada vez más elevada de los ciudadanos hacia las instituciones públicas y privadas.

Cuadro 4. Principales etapas en la expansión de la evaluación. Cuarta ola: finales 2000 – actualidad (2014) (cont.)

- En el ámbito OCDE, se plantea con cierta fuerza el modelo llamado neo-weberiano (NW), estructurado alrededor de unos principios de Administración Pública que guían la reforma de las administraciones (Pollitt y Bouchaert, 2004; SIGMA). Se reafirma el rol del Estado como facilitador principal de soluciones a los nuevos problemas planteados por la globalización; el rol de la democracia representativa como elemento legitimador del aparato del Estado; el rol del derecho administrativo, revisado y modernizado, en la preservación de los principios que regulan la relación entre el Estado y los ciudadanos.
- Otras tendencias, como el Public Value Governance (PVG) (Bryson, Crosby, Bloomberg, 2014; también Chelimsky 2014), subrayan el predominio de valores democráticos más allá de los propios de efectividad y de eficiencia, y el rol especial de los gobiernos como garantes de estos valores y como generadores de valor público, equivalente a los efectos finales (o outcomes) conseguidos en relación con el interés general. El valor público emerge a partir de procesos inclusivos de diálogo y deliberación, establecidos entre miembros de sectores diferentes de la comunidad⁹. Los ciudadanos pasan de ser administrados, votantes, clientes o beneficiarios de la acción de los gobiernos a participar activamente en la resolución de problemas de interés público y en la co-creación de valor público.
- Por otra parte, en esta etapa se impulsa al desarrollo de sistemas de información pública estadística y administrativa, de la mano de las nuevas tecnologías, así como de las políticas de transparencia y de las leyes de acceso a información pública y de reutilización de información administrativa.
- Asociado a lo anterior, auge del análisis cuantitativo de datos anonimizados de registros públicos accesibles a todo tipo de usuarios ("Big Data").

Características de la evaluación

- Vuelta a la concepción cientifista-experimental de la evaluación: Evidence-based policy making, aunque aún lejos de ser el paradigma dominante. Esa orientación ha empezado a contagiar también la evaluación de programas financiados por fondos estructurales de la Unión Europea, en particular en el ámbito de las políticas activas de empleo y las políticas sociales (European Commission, 2014).
- Al mismo tiempo, la evaluación se extiende a otros campos de conocimiento científico, como el medio ambiente, más allá de las ciencias sociales y de las ciencias de la salud, que han constituido sus feudos tradicionales.

9. "Public values and public value are not the exclusive province of government, nor is government the only set of institutions having public value obligations, [though clearly] government has a special role as guarantor of public values." (Beck Jørgensen i Bozeman, 2007, pág. 373-374).

Cuadro 4. Principales etapas en la expansión de la evaluación. Cuarta ola: finales 2000 – actualidad (2014) (cont.)

- La evaluación rigurosa de impacto adopta aproximaciones “mixed methods”, aplicando técnicas de análisis cualitativo para tratar de comprender y explicar mejor procesos e impactos.
- Cierta saturación institucional en algunos casos (“inflación de la evaluación”) sin rendimientos claros.
- Irrupción del concepto y prácticas de “innovación social” y de “inversión social”, vinculadas a la práctica de la evaluación de impacto.
- Colaboración con el sector privado, que a través de “partenariados público-privados” aporta fondos para testar de forma experimental intervenciones innovadoras del sector público o del tercer sector, ya sea como actividad filantrópica o como inversión.
- Ola de “institucionalización flexible” de la evaluación en países anglosajones fuera de la UE, en particular en Estados Unidos ¹⁰.
- Emergencia de los “sistemas de evaluación” y las “políticas de evaluación” en la literatura internacional. Creciente interés por el estudio de las pautas de institucionalización y de sus resultados.

Fuente: elaboración propia.

10. Ver, por ejemplo, los siguientes memorándums sobre evaluación elaborados por la Office of Management and Budget (OMB):

- Increased Emphasis on Program Evaluations (<http://1.usa.gov/bRGwej>), de Octubre de 2009.
- Evaluating Programs for Efficacy and Cost-Efficiency (<http://1.usa.gov/aC2HSS>), de Julio de 2010.
- Use of Evidence and Evaluation in the 2014 Budget, (<http://1.usa.gov/KZ9fsR>), de Mayo de 2012.

III. La institucionalización de la evaluación en Europa

1. La evaluación en los países nórdicos: el caso de Suecia.

1.1. Rasgos generales de la evolución de la evaluación en los países nórdicos.

La evaluación en los países nórdicos, y especialmente en Suecia y en Dinamarca, se ha desarrollado en estrecha relación con el establecimiento y expansión del Estado del Bienestar. Su institucionalización ha seguido fases y trayectorias diferentes, y su grado de desarrollo varía entre países.

Se dice que Suecia formó parte de la primera oleada de institucionalización de la evaluación, mientras que Dinamarca, Finlandia y Noruega formaron parte de la segunda oleada (Furubo y Sandahl, citados por Forss y Rebien, 2014). También han seguido diferentes trayectorias tanto en la evolución del sector público como en la profesionalización de la evaluación (Holm, 2006), lo que ha afectado de manera diferente a la demanda y a la oferta de evaluaciones, así como su utilización. Sí parece un rasgo común el hecho de que está mucho más presente en el sector público que como disciplina académica singularizada:

"...evaluation as a separate field of research and study hardly exists." (Holm, 2006, pág. 141)

Otro rasgo común parece ser que, dentro del sector público, los tribunales de cuentas juegan un papel importante en el campo de la evaluación, aunque no propiamente en lo que se refiere a la evaluación de efectividad o impacto de las políticas públicas, sino en la variante de evaluación del rendimiento o *Performance Evaluation* del ejecutivo y no sólo en la auditoría de las cuentas públicas:

"In all the Nordic countries auditing of state accounts is mentioned when discussing evaluation." (Holm, op. cit., pág. 141)

En el Ejecutivo, en general, las evaluaciones se llevan a cabo de forma descentralizada por los diferentes ministerios sectoriales, y su práctica está más arraigada en ciertos

ámbitos de intervención pública, como es el caso de las políticas educativas, sociales, medioambientales y de ayuda al desarrollo.

Por otra parte, la entrada de Suecia, Finlandia y Dinamarca en la UE promovió también la práctica de la evaluación, como en el resto de países miembros.

En cuanto a posibles elementos de una eventual “cultura” o “tradición” común de evaluación, Forss y Rebien (*op. cit.*) señalan, a modo de hipótesis pendientes de contraste, las siguientes:

- El hecho de que la función del Defensor del Pueblo u *Ombudsman* sea originaria de los países nórdicos indica una fuerte tradición de evaluación presente en el sistema de control público y parlamentario de la Administración y del poder ejecutivo.
- La tradición de audiencias públicas de proyectos de ley –protegida por la Constitución en los países nórdicos– proporciona una cultura de implicación de partes interesadas o *stakeholders*, vinculada a la evaluación *ex ante*. Esta tradición de evaluación *ex ante* va de la mano con una fuerte tradición de evaluación *ex post* para ajustar y aprobar nueva legislación.
- El sistema de bienestar financiado con fondos públicos –gracias a uno de los sistemas impositivos con tipos más elevados del entorno internacional– da pie a un papel destacado de la evaluación para asegurar un sector público correctamente gestionado.

Por lo demás, los países nórdicos coinciden también en contar con poblaciones pequeñas en términos cuantitativos, en haber emprendido procesos de modernización bastante tarde, situarse en la periferia de Europa, contar con estados del bienestar y sectores públicos de gran alcance y ser también sistemas democráticos multipartidistas (Forss y Rebien, *op. cit.*).

Bohni Nielsen y Winther (2014) comparan la práctica de la evaluación en los países nórdicos e identifican un cierto “dominio” de Suecia, que podría obedecer a los factores siguientes:

- En primer lugar, la tradición sueca de evaluación evolucionó más rápidamente y dispone de una mayor masa crítica en términos de producción científica y calidad de ésta.
- En segundo lugar, la institucionalización de la evaluación en Suecia (y en Finlandia) es hasta cierto punto diferente de la pauta seguida en Noruega y Dinamarca. En concreto, debido a la existencia en los primeros de unidades internas de evaluación en organismos del Gobierno y a la existencia de comisiones nacionales con responsabilidades de evaluación. Así, en Suecia y Finlandia, esas comisiones y unidades de evaluación tienden a encargar evaluaciones a expertos académicos, mientras que en Dinamarca y Noruega hay una práctica evaluadora más pluralista, con más participación de consultoras privadas pero también de ONG y otras entidades del tercer sector.

1.2. La evaluación en Suecia

El origen de la evaluación en Suecia se remonta a los años 30, aunque de forma más sólida a los años 50, centrada fundamentalmente en políticas de reforma educativa impulsadas en esa época (Foss Hansen, citado por Holm, *op. cit.*). Este desarrollo ha venido determinado por una cultura política caracterizada por una larga tradición (la más antigua de Europa) de libertad de acceso a información pública, así como por un sistema político que da gran importancia al diálogo y al consenso (Kuhlmann, 2010).

Según Furubo y Sandahl (2002), dos factores destacan especialmente como determinantes del desarrollo de la evaluación en Suecia. En primer lugar, la práctica, muy arraigada, de crear comisiones *ad hoc* para analizar y decidir sobre determinadas políticas públicas. La evaluación ha constituido tradicionalmente un *input* relevante en el trabajo de estos órganos y en la toma de decisiones a la que dan apoyo. En segundo lugar, la elevada descentralización del gobierno sueco, y la distinción conceptual y organizativa entre ministerios -13 actualmente, que emplean 5.000 servidores públicos- y agencias -alrededor de 250, con 200.000 empleados- (ESV, 2013).

Esa distinción refleja el principio de separación entre la formulación de políticas públicas, que es competencia del ejecutivo, y su administración y ejecución, que es responsabilidad de agencias que mantienen un elevado grado de autonomía respecto de los ministerios, los cuales no pueden ejercer un control detallado de sus actividades. De hecho, el presupuesto de las agencias es independiente del de los ministerios sectoriales, y es controlado directamente por el Ministerio de Finanzas. En cambio, las agencias deben rendir cuentas de sus actividades y resultados ante el Gobierno como órgano colegiado, lo que ha facilitado una amplia penetración de la práctica evaluadora y de la utilización de sus resultados.

Durante los años 80 y 90, el desarrollo de actividades de evaluación en Suecia se vio influido por las sucesivas crisis económica y financiera, así como por los procesos de reforma del sector público y de la administración que tuvieron lugar en esa etapa. Desde principios de los años 80, cuando se enfrentó a una importante crisis económica que puso en jaque su tradicional estado del bienestar, Suecia se embarcó en un doble proceso de modernización. De un lado, en 1985 el Gobierno sueco emprendió un proceso de profundización de la descentralización territorial, ya bastante amplia comparada con otros países europeos. Además, entre 1988 y 1993 se emprendieron una serie de reformas en el ámbito de la gestión de las finanzas públicas, incluyendo la introducción del presupuesto por resultados, en el marco de ola de *New Public Management*. Así, la línea de reforma “tradicional”, orientada a una mayor descentralización del gobierno, con la consiguiente reducción de competencias y recursos de las agencias ejecutivas estatales a favor de los entes locales, se combinó con la introducción de medidas típicas del NPM en el gobierno central, que afectaron en especial a la relación entre éste y las agencias ejecutivas, orientándolas a la gestión por resultados. Esos

mismos sistemas de gestión fueron introducidos también en los municipios, como veremos más adelante.

Por otra parte, a principios de los 90 Suecia sufrió una dura crisis financiera, que situó su déficit presupuestario como uno de los más elevados de Europa (cerca del 11% del PIB en 1993) e hizo aumentar su deuda pública hasta el 70% del PIB (ESV, 2013). Esa crisis llevó a una serie de reformas en el marco institucional de gestión y control de las finanzas públicas que han continuado hasta la actualidad, y que supusieron la puesta en práctica de sofisticados sistemas de seguimiento y análisis de la gestión y del uso dado a los recursos presupuestarios.

La institucionalización de la evaluación en el gobierno central

La práctica de la evaluación en las instituciones del gobierno central está ampliamente extendida y también diversificada en cuanto a sus formas de provisión, hasta el punto de que algunos autores se refieren a múltiples “sistemas de evaluación” que se han ido constituyendo a través de los años en el interior de instituciones públicas en los ámbitos de la educación, las políticas sociales, las políticas medioambientales, la ayuda al desarrollo, etc. (Holm, *op. cit.*).

Encontramos, por un lado el sistema de comisiones creadas *ad hoc* mencionadas anteriormente, articuladas cada una de ellas alrededor una determinada política pública. La composición de esas comisiones, aunque variable en cada caso, suele incluir diferentes “partes interesadas”: desde representantes de partidos políticos hasta sindicatos y patronales, grupos de interés, asociaciones, etc., en función de ámbito de que se trate, así como expertos académicos en la materia y profesionales del sector público. Todos los miembros tienen acceso a la misma información que va a utilizarse en el análisis y -lo que probablemente es más importante y denota el grado de influencia y de implicación en las tareas de análisis- participa en la formulación de las preguntas clave que van a guiar los trabajos de la comisión, y que a menudo se refieren a aspecto de efectividad y coste-efectividad de las políticas.

Por la naturaleza de sus trabajos y por el tipo de material en que se basan, esas comisiones pueden considerarse como el principal canal de transmisión en Suecia entre el conocimiento académico existente sobre diferentes ámbitos de política pública y el proceso de toma de decisiones al respecto en las instituciones públicas. Según Furubo y Sandahl (*op. cit.*), aunque no todas las comisiones cuentan con expertos académicos entre sus miembros, sí es muy habitual que se recurra a material académico como base de sus trabajos.

Por otra parte, las agencias ejecutivas han asumido también funciones de evaluación. En una primera etapa, estas actividades se desarrollaron especialmente en el ámbito educativo, por parte de la *Swedish Board of Education*, en el de la energía y también en la cooperación al desarrollo, gestionada por la *Swedish International Development Agency* (SIDA), y posteriormente se extendieron a otras áreas de gobierno.

Sin embargo, no existen estudios empíricos recientes que permitan establecer el grado de cobertura y la regularidad en la práctica de la evaluación en las comisiones *ad hoc* ni en las agencias ejecutivas¹¹.

Las actividades de evaluación desarrolladas por el ejecutivo son complementadas a su vez por las que emprenden organismos independientes responsables de la evaluación externa de las políticas gubernamentales, como la *Swedish National Audit Office*¹², que lleva a cabo diversos informes de *Performance Auditing* cada año, centrados en analizar la eficiencia de las actuaciones del gobierno:

*"The starting point here is how well the subject of the audit is achieving its goals and the appropriateness of the organisation, operations, process, or function for the purpose. The goal of the audit is to make sure that the state, with regards to the interests of the general public, is getting good results for its efforts."*¹³

En el web de la institución se publican cada año numerosos informes, orientados muchos de ellos a análisis de efectividad¹⁴.

Otro actor importante es la *Swedish Financial Management Authority (Ekonomistyrningsverket - ESV)*¹⁵. La ESV es una agencia del Gobierno central, adscrita al Ministerio de Finanzas, responsable de la gestión eficiente del presupuesto. Sus funciones incluyen también la provisión de herramientas a las agencias en material de gestión por resultados, auditoría de fondos de la UE, etc., además de análisis *ex post* y también prospectivo de las finanzas públicas.

Por otra parte, el Ministerio de Finanzas cuenta también con otras agencias que aportan diferentes insumos de información y análisis sobre el uso dado a los recursos públicos, entre los que destacan la agencia estatal de estadística (*Swedish Office of Statistics - Statistiska Centralbyrån*)¹⁶ y el *National Institut for Economic Research (NIER)*¹⁷.

Además, el Parlamento puede establecer comisiones específicas de evaluación sobre determinadas intervenciones públicas.

11. Una visita a la sección de publicaciones del web del Gobierno sueco (<http://www.government.se/sb/d/574>, el 1 de febrero de 2015) permite identificar tan sólo dos documentos que contienen en su título la palabra "evaluación"; ninguno de ellos es reciente y sólo uno se refiere a una política pública: *Midterm evaluation of the Swedish Rural Development Programme 2007-2013*. Ministry of Enterprise and Innovation, Reports. 16 November 2010. El otro es: *Evaluation of the Group Proceedings Act - Summary in English* Ministry of Justice, Information material. 28 October 2008. Sin embargo, es posible que la versión en sueco del web contenga más publicaciones.

12. <http://www.riksrevisionen.se/en/Start/About-us/Our-fields-of-operation/>.

13. Web de la SNAO visitada el 1 de febrero de 2015.

14. **Agunos ejemplos correspondientes a 2014:** *Primary health care governance – based on need or demand?*; *Food control – is central government fulfilling its responsibility?*; *Utilising and developing the competencies of newly arrived immigrants – the right measure at the right time?*; *Effects of amended rules for the part-time unemployed*.

15. <http://www.esv.se/ESV/English/>

16. <http://www.scb.se/en/>

17. <http://www.konj.se/english>

En cualquier caso, hoy en día la evaluación es un componente integral en áreas como el presupuesto, la educación superior y la política medioambiental.

Prácticas de evaluación en los gobiernos locales

La penetración de la evaluación en los gobiernos locales suecos ha seguido una pauta “de abajo arriba”, en la que la *Swedish Association of Local Authorities (Svenska Kommunförbund)* ha jugado un rol muy importante (Kuhlmann, *op. cit.*; Wollmann, *op. cit.*). La disponibilidad y calidad de la información estadística municipal en la *Swedish Office of Statistics* desde finales de los 80 ha facilitado enormemente la extensión de estas prácticas. Esa información incluye numerosos indicadores y datos sobre diversos servicios gestionados a nivel local y sus costes (escuelas, servicios sociales para personas mayores y para personas con discapacidad, etc.).

La riqueza de información, la implicación voluntaria de las entidades locales y la coordinación de la asociación de municipios, ha facilitado no sólo el análisis de cada municipio, sino también la comparación entre municipios. Además, con el paso de los años los indicadores contemplados se han ido enriqueciendo y extendiendo a aspectos de resultados e impactos de las políticas municipales, en contraste con una primera fase en que se centraban fundamentalmente en aspectos más “convencionales”, relativos a costes y satisfacción de los usuarios (Wollmann, *op. cit.*). La accesibilidad a esa información por parte del público afecta positivamente, según algunos autores, a su utilización en la toma de decisiones por parte de las autoridades locales:

“As performance indicators are accessible to all, they are used not only for the internal steering of the authority and the ‘fiscal control of the businesses’. They are also taken seriously by politicians, particularly during elections.” (Kuhlmann, *op. cit.*, pág.343).

La profesionalización de la Evaluación

Suecia cuenta con una red profesional de evaluación que se puso en marcha en 1993 y se refundó posteriormente en 2003 bajo el nombre de *Swedish Evaluation Association*. De forma similar a otras redes profesionales existentes en otros países y en el entorno profesional, la finalidad de la asociación sueca es mantener un debate sobre el papel de la evaluación y de los evaluadores en la sociedad y en las instituciones públicas, desarrollar la competencia profesional de los evaluadores a través de guías y recursos, formación, etc., y establecer vínculos con otros organismos afines tanto en el ámbito nacional como internacional.

El sistema de evaluación en Suecia

Valoración general

En Suecia existe un sistema de evaluación consolidado, tanto en la administración central como en la local. Ese sistema ha emergido gracias a la larga tradición sueca de transparencia y rendición

de cuentas, al gran avance experimentado por las ciencias sociales ya desde el período de entreguerras, siguiendo la estela de Estados Unidos, a la calidad de la democracia sueca y al desarrollo de su Estado del Bienestar. En ese contexto, y dadas las peculiaridades del entorno institucional público y, en particular, de la configuración del Ejecutivo, el sistema de evaluación es pluralista y descentralizado, aunque instancias como la Swedish Financial Management Authority ejercen una cierta coordinación del sistema. Del mismo modo, agencias como la Swedish Office of Statistics -Statistika Centralbyrån y el National Institut for Economic Research (NIER) aportan know-how y recursos de información. Aunque para la confección de este informe no se ha podido analizar directamente una muestra de evaluaciones recientes, de la literatura y documentos consultados se desprende la existencia de múltiples enfoques de evaluación, incluyendo evaluaciones rigurosas de impacto. Si bien la difusión pública de las evaluaciones es una práctica generalizada, no hemos tenido acceso a datos que nos permitan valorar el grado de utilización efectiva de los resultados en la mejora de los programas evaluados.

2. La evaluación en los Países Bajos

En el caso de Holanda, Frans L. Leeuw (2009) establece el inicio de una política explícita de evaluación en el gobierno central en la década de los 90. Sin embargo, la actividad evaluadora ya había hecho acto de presencia en el período de entreguerras y, posteriormente en los años 60 y 70, y se había ido desarrollando a finales de esa década y a lo largo de los 80.

Su origen se sitúa en la rica tradición del país en investigación social aplicada a las políticas públicas, y su avance a partir de finales de los 70 se explica fundamentalmente por presiones presupuestarias derivadas de la importante crisis económica de mediados de esa década, que agudizaron el interés por conocer el impacto de los programas públicos. Durante los 80 su eclosión coincide con introducción del *New Public Management* en el gobierno y administración holandeses, junto con una voluntad reforzada de control parlamentario sobre el ejecutivo y el inicio de auditorías de resultados (*performance audits*) por parte de la *Netherlands Court of Audit* (NCA), que pusieron en evidencia la falta de información sobre efectividad y costes de las intervenciones públicas. Leeuw destaca también como palanca de impulso una cultura administrativa favorable, excepcionalmente abierta a la innovación y en la cual la evaluación fue acogida como un útil instrumento de aprendizaje.

Los orígenes

El origen de la evaluación de políticas públicas en los Países Bajos se remonta a los años 20, en los que se desarrollan, en el seno de la comunidad científica, líneas de investigación sobre las políticas sociales, y en particular sobre desarrollo rural y urbano. Durante los años 50 y 60, ciertos investigadores de ese campo se interesan

por los programas públicos, y sus investigaciones son encargadas y financiadas en gran medida por el Gobierno, sin que tengan ninguna relación con las funciones de auditoría, control o análisis presupuestario-contable que son llevadas a cabo por los diferentes departamentos del gobierno como parte de sus funciones ordinarias.

Los estudios de investigación encargados y dirigidos a responsables públicos en esa etapa, tienen por finalidad responder preguntas sobre las causas de determinadas problemáticas sociales y sus posibles remedios, pero se trata de estudios aislados, sin ninguna perspectiva transversal y con una débil utilización de resultados por parte de los responsables políticos (Jacob, 2005).

En los años 70, las presiones impuestas por la crisis económica internacional sobre la hacienda pública forzarán un desarrollo de la evaluación vinculado al control del déficit presupuestario (Varone y Jacob, 2004). En efecto, a principios de los años 70 el Ministerio de Finanzas pone en marcha un comité interdepartamental para el desarrollo del análisis de las políticas (*Commissie voor de Ontwikkeling van Beleidsanalyse* – COBA), y le encarga promover el análisis y la evaluación de las políticas públicas. Ese comité, en el que están representados todos los departamentos del Gobierno, así como la Oficina del Plan y la Oficina Central de Estadística, es asistido por un equipo reducido, con sede en el Ministerio de Finanzas y, entre otras actuaciones, publica una revista trimestral de análisis de políticas (*Belidsanalyse*), que incluye experiencias de los Países Bajos y también del extranjero.

Esa iniciativa se inscribe en el movimiento de Planificación, Programación y Presupuestación (*Planning, Programming and Budgeting System*) y de una gestión por objetivos (*Management by objectives*), y tiene por objetivo proporcionar una base institucional para la utilización del análisis y de la evaluación de las políticas públicas en el proceso presupuestario. Se trata de tener en cuenta en la gestión y presentación del presupuesto los medios, los instrumentos, los objetivos y los efectos de las políticas.

En 1979, el comité presenta un balance en el que destaca la resistencia política y administrativa en el avance hacia sus metas. En particular, las dificultades derivadas de un sistema multipartito que dificulta la definición de objetivos claros; la débil implicación de los funcionarios en un proceso de reflexión sobre su propio trabajo y también problemas de tipo metodológico. Esas dificultades parecen determinar la finalización de los trabajos y de la propia experiencia de la COBA en 1983.

Sin embargo, esa línea de actuación se prolonga en el seno del Ministerio de Finanzas, que pone en marcha un Departamento de Análisis de Política (*Afdeling Beleidsanalyse*). Así, aunque la experiencia COBA acaba en fracaso, permite introducir la evaluación en la agenda política del ejecutivo.

A principios de los años 80, las reflexiones sobre la institucionalización de la evaluación se multiplican. Un informe publicado por la Comisión parlamentaria de asuntos públicos (*Commissie Hoofdstructuur Rijksdienst*) preconiza el recurso a normas experimentales a fin de permitir el desarrollo de la práctica de la evaluación. Esa Comisión organiza una encuesta entre cincuenta y siete ministros, altos funcionarios y parlamentarios para analizar la situación de la evaluación. Los resultados muestran que se llevan a cabo pocas evaluaciones, que su calidad es insuficiente y que los resultados no se explotan suficientemente. Por otra parte, más del 80% de los entrevistados considera que es necesario institucionalizar la evaluación de las políticas públicas (Jacob, *op. cit.*).

En 1982, entre las soluciones preconizadas y puestas en práctica para controlar el déficit público figura el “procedimiento de reconsideración”. Ese procedimiento se interesa por los efectos de las políticas pero busca, sobre todo, servir de orientación para la realización de recortes presupuestarios. Sea como sea, estimula una fuerte demanda de evaluaciones: 280 estudios de reconsideración son llevados a cabo por comités que se benefician de una independencia de actuación y de juicio garantizada por un acuerdo entre los miembros del Gobierno, que establece que los ministros se abstendrán de interferir en sus trabajos. Desde una perspectiva comparada, se toma como referencia el ejemplo de la *National Audit Office* del Reino Unido y la *General Accounting Office* de los Estados Unidos.

Por otra parte, el procedimiento de reconsideración permite de nuevo establecer un puente entre la comunidad académica y la administración. Esta vez se persigue la implicación de altos funcionarios y no solo de unidades operativas, como en los primeros años de introducción de la evaluación.

La política de evaluación en el ejecutivo

La política de evaluación propiamente dicha inició su formalización en 1991, a partir de una nueva legislación sobre control gubernamental aprobada en ese año (la *Government Accounting Act*), que establecía la obligación de los ministerios de evaluar sus políticas cada cinco años. Y a través de un breve documento emitido por el gobierno –el *Cabinet Memorandum on Policy Evaluation*– que establecía unos principios generales rectores de la práctica evaluadora. Se descartó, por tanto, una regulación detallada de la evaluación, que por otra parte ya se mencionaba en la ley de finanzas y en otras disposiciones administrativas.

El memorándum establecía, entre otras cuestiones, la necesidad de realizar planes de evaluación por parte de departamentos y organismos y de vincular la información resultante a la toma de decisiones y al presupuesto. La concreción de métodos, estándares y cuestiones relativas al uso de los resultados de las evaluaciones se dejaban en manos de los diferentes departamentos y agencias, aunque por otro lado se favoreció

la creación de unidades de evaluación que incorporaron esas funciones. Así, ese nuevo marco normativo permite desarrollar y generalizar la práctica de la evaluación en un gran número de organismos públicos, sobre la base de estándares y procedimientos diseñados de forma homogénea.

Otros hitos en el desarrollo de la evaluación son la Ley del Gobierno (1990), que enuncia un principio general sobre la evaluación, y la Ley del procedimiento presupuestario (1998) (*Van Beleidsbegroting tot beleidsverantwoording* – de la política presupuestaria a la política de la responsabilidad), que pretende vincular la presentación del presupuesto a los resultados de la actuación pública, a través de la formulación de preguntas como “¿Cuáles son los objetivos? ¿Cómo vamos a proceder para alcanzarlos? ¿Cuáles son los costes implicados? ¿Hemos obtenido lo que buscábamos?” (Jacob, *op. cit.*). Ese procedimiento implicaba, además, la obligación de cada ministro de presentar un informe anual describiendo los objetivos y los logros obtenidos. La orientación general era no tanto sancionadora como de aprendizaje basado en la medida de los resultados.

Se trata de un proceso progresivo y de acumulación de experiencia en el seno de los ministerios del gobierno holandés, que lleva más adelante, a partir de 2002, a la adopción de cláusulas de evaluación de carácter general (ej. el Reglamento sobre el análisis del rendimiento y la evaluación, o *Regeling Prestatiegegevens en Evaluatieonderzoek*). Esas cláusulas permitirán avanzar hacia una aplicación más o menos uniforme, estandarizada, de la práctica evaluativa en todos los ministerios.

Por otra parte, de la mano del *New Public Management*, a lo largo de los 90 se impulsaron también los sistemas de indicadores de seguimiento / resultados (*performance indicators*) en diferentes áreas de gobierno, asociados, como parte del paquete de reformas, a la creación de agencias departamentales ya iniciada en los 80 y que en los 90 tomó un nuevo impulso.

Posteriormente, ya en 2001 -año de creación de la *Dutch Evaluation Society*-, se reforma la ley de finanzas (*Budget and Accounts Act*), reforzándose formalmente el papel de la evaluación. El nuevo marco normativo establece una articulación más clara de la función evaluadora en el gobierno holandés, en particular de aspectos relacionados con el uso de diferentes enfoques de evaluación, la frecuencia de su realización, cuestiones metodológicas –refinando, por ejemplo, los criterios relativos a evaluaciones de impacto-, y otros relativos a la utilización y a la difusión de las evaluaciones. De nuevo, los departamentos y agencias ejecutivas son responsables de concretar y desarrollar esos aspectos en cada caso. Para ello, cuentan normalmente con unidades internas, tanto de inspección como de finanzas, que llevan a cabo actuaciones de evaluación (Wollmann, *op. cit.*).

Según Leeuw (*op. cit.*) se constata un avance general, aunque se da un cierto solapamiento de funciones entre organismos públicos de auditoría, inspección y evaluación.

Esta situación y la sobrecarga administrativa que conlleva, junto con la percepción pública de focos de ineficiencia persistentes, ha llevado al gobierno, desde finales de la década de 2000, a plantear una revisión del modelo actual de resultado todavía incierto.

En general, la institucionalización de la evaluación de políticas públicas en el ejecutivo y en la Administración holandesa se caracteriza actualmente por la búsqueda de soluciones prácticas en el seno de los diferentes departamentos, que se benefician del apoyo de órganos centrales que supervisan el desarrollo de la evaluación. El Ministerio de Finanzas estimula y coordina el dispositivo, especialmente a través de la puesta en marcha de una red en la cual se pone el acento en la relación entre la evaluación y el presupuesto por programas. El Ministerio de Finanzas no se propone tomar el control de ese dispositivo, ya que entiende que la evaluación es responsabilidad de cada ministerio, si bien puede llevar a cabo, igual que la *Netherlands Court of Audit*, meta-evaluaciones y otras actuaciones para supervisar y comprobar cómo se está desarrollando la evaluación. Su rol consiste fundamentalmente en una mezcla de incitación, de control y de sanción simbólica, llamando al orden en caso necesario (Jacob, *op. cit.*).

El rol del Parlamento y de la *Netherlands Court of Audit*

El Parlamento holandés empieza a implicarse en la evaluación de políticas públicas a principios de los años 80, a partir de una investigación sobre política industrial que llega a conclusiones muy críticas sobre la actuación del Gobierno. En definitiva, esas conclusiones ponen de relieve que el enorme esfuerzo presupuestario dedicado a ese ámbito de intervención no ha obtenido resultados. Esa crisis abre una ventana de oportunidad que no es, sin embargo, aprovechada por el Parlamento para crear una instancia interna de evaluación. Ello se explica, por una parte, porque en el país las relaciones entre Parlamento y Ejecutivo reposan sobre la noción de responsabilidad del segundo, que acomete él mismo el impulso de la evaluación de sus actividades. Y por otra, por la total independencia del Tribunal de Cuentas o *Netherlands Court of Audit* tanto respecto del Parlamento como del Ejecutivo (Jacob, *op. cit.*).

De hecho, a raíz de esa situación y de la constatación más general de la falta de seguimiento sistemático sobre la efectividad de las intervenciones públicas, la NCA inicia *Government wide audits* sobre la preparación del presupuesto, la implementación de las políticas y la utilización de las subvenciones. La finalidad de esas auditorías es poner en evidencia la base de conocimiento de que disponen los departamentos del gobierno para actuar con eficacia y eficiencia. Las conclusiones son claras sobre la escasez del conocimiento disponible.

Sobre el papel de la *Netherlands Court of Audit* en la institucionalización de la evaluación en Holanda, en este país, como en el caso de Suiza, la evaluación forma parte del núcleo de actividad de este organismo, a diferencia de lo que ocurre en Francia y en

Bélgica, donde han permanecido en un rol más tradicional, de control de conformidad y legalidad (Jacob. *op.cit.*). En los Países Bajos, la NCA adopta una posición clara en materia de evaluación y se implica a fondo en el acompañamiento de las actuaciones realizadas y en el establecimiento de los dispositivos de evaluación. De hecho, según Jacob (*op. cit.*), dedica tantos esfuerzos a la evaluación como a sus funciones tradicionales de control.

Desde 1991, a raíz de la adopción de la *Dutch Government Accounting Act*, la misión de la NCA se ve modificada, poniéndose el acento en la eficacia y la eficiencia. Ello se traduce en la realización de *performance audits* por parte de este organismo, definidas como verificaciones tendentes a evaluar de qué forma la entidad objeto de control dedica los recursos públicos que le son asignados a la consecución de la misión que tiene encomendada, en el marco de los principios de economía, eficiencia y eficacia.

Desde finales de los 80, el Tribunal se ha dotado también de la misión de controlar la calidad de las evaluaciones producidas por los ministerios, a través de metaevaluaciones. Esos ejercicios tienen una finalidad formativa, de ayuda a los departamentos ministeriales para que desarrollen sus estructuras y sus prácticas evaluadoras. Por otra parte, se asume un desarrollo desigual de la evaluación entre los diferentes departamentos del Gobierno, y en los casos más débiles, la NCA puede asumir directamente la realización de *performance audits* buscando un efecto “demostración”, pedagógico. En la literatura sobre el desarrollo de la evaluación en los Países Bajos, se encuentran referencias a estudios de ese tipo llevados a cabo por la NCA en 1990 y en 1998 (Bemelmans-Vidéc, 2002; Jacob, *op. cit.*; Leeuw, *op. cit.*), en los que se detecta una gran disparidad, tanto en lo que se refiere al número de evaluaciones realizadas por los diferentes ministerios, como a su relevancia o al presupuesto empleado. Entre las posibles causas de esa situación, se identifica la falta de una definición clara de lo que debe ser evaluado cada cinco años. A partir de esos estudios, el Ministerio de Finanzas emprende revisiones sucesivas de la política de evaluación ayudado por la NCA, que mantiene un rol de “help desk” hacia los departamentos en este ámbito, y también de seguimiento del uso de las evaluaciones. Esta utilización parece elevada, aunque las recomendaciones más seguidas son aquellas que afectan a aspectos organizativos, más que substantivos de las políticas (Jacob, *op. cit.*).

Más recientemente, en 2012, la NCA realizó un nuevo informe sobre la práctica de la evaluación en el ejecutivo, referido al período 2006-2010 (NCA, 2012). Según ese informe, en el período de referencia se llevaron a cabo 1.200 evaluaciones de políticas públicas, de las cuales 350 (29%) analizaron la efectividad de las políticas en términos de impacto. Los programas objeto de evaluación sumaron un presupuesto de 51 mil millones de euros, lo que supone el 46% del gasto presupuestario en 2010. El estudio no entró a analizar la calidad de las evaluaciones, pero sí detectó errores a la hora de clasificar como evaluaciones de impacto estudios que en realidad no lo son, así como un

déficit de información al Parlamento sobre las evaluaciones que se realizan y también, especialmente, sobre las razones para no evaluar la efectividad de ciertas políticas.

En un informe posterior, el Tribunal de Cuentas holandés no detecta apenas mejoras, confirmándose la falta de evaluaciones sobre la efectividad de cerca de la mitad de las políticas públicas financiadas por el Gobierno central:

“There has been no change in the general picture that emerged from our initial audit. The effectiveness of slightly less than half of the policies funded by central government was not evaluated in the past six years... Opportunities are therefore being missed to apply the potential lessons learned in the formulation of new policy. It is important that the government has information on the effectiveness of policy in as many areas as possible. Public money will then not be invested in policy that does not work effectively, if at all, and schemes that are effective will not be cut.” (NCA, 2013)

Para poner remedio a la situación la NCA insta a los departamentos del gobierno a elaborar programas de evaluación que establezcan de forma razonada qué evaluaciones y de qué tipo deben abordarse:

“We believe spending cuts and reforms should be implemented in a responsible manner based on reasoned choices. To this end, the government must have the relevant information it needs to determine the impact of spending cuts – not only budgetary information but also social information... The objective need not be to evaluate the effectiveness of all policy at all times, but every ministry should have a reasoned evaluation program laying down what kinds of evaluation should be carried out when and what kinds have the most added value.” (NCA, 2013)

No se han publicado, en el momento de redactar este informe, estudios ulteriores que permitan conocer el grado de seguimiento de esas recomendaciones.

Por otra parte, en lo que respecta al desarrollo de la evaluación en el ámbito local, según algunos autores los Países Bajos presentan un cierto retraso respecto a otros países como Suiza o Francia (Jacob, 2005), aunque se ha comenzado a desarrollar a partir de 2002. Sin embargo, la falta de estudios empíricos al respecto imposibilita profundizar más en este ámbito a la hora de redactar este informe.

La profesionalización de la evaluación

En los Países Bajos, la comunidad evaluadora ha carecido durante largo tiempo de una asociación profesional de evaluación, a pesar de ser uno de los países fundadores de la Sociedad Europea de Evaluación, especialmente a partir del impulso dado al proyecto

por parte del equipo directivo de la NCA, y del apoyo económico del Ministerio de Finanzas holandés (Jacob, *op. cit.*). Posteriormente, en 2002, se crea una asociación nacional, la *Beroepsvereniging (VIDE)*¹⁸ que agrupa inspectores, auditores y evaluadores, con la finalidad de promover un intercambio de conocimiento a nivel nacional que mejore la práctica evaluadora.

Sin embargo, a pesar de la puesta en marcha relativamente tardía de la VIDE, en los Países Bajos el reconocimiento de la evaluación como disciplina académica y profesional empieza ya antes de la década de los 90 (Bemelmans-Videc, 2002), y actualmente hay gran número de programas universitarios, tanto de grado como de posgrado, que la incluyen. También está muy presente en la producción científica y en el sector privado de consultoría, existiendo un mercado muy activo de estudios de evaluación.

El sistema de evaluación en los Países Bajos

Valoración general

El desarrollo de la evaluación en los Países Bajos se enmarca en una rica tradición científica y una cultura administrativa muy receptiva a la innovación y a la mejora continua. Después de una primera etapa de evaluaciones puntuales y externas, empieza una aproximación más sistémica, a raíz de la crisis económica de mediados de los 70 y la adopción del presupuesto por programas y de la gestión por objetivos. Esos primeros pasos se verán complementados por el refuerzo de las tareas de control del ejecutivo por parte del Parlamento y de la Netherlands Audit Office en los años 80. El sistema de evaluación se desarrolla propiamente en la década de los 90, a través de regulaciones sucesivas que establecen la obligación de evaluar, definen qué debe entenderse por evaluación –contemplando diferentes enfoques–, tratan de ordenar la práctica evaluadora a través de la confección de agendas de evaluación, así como de establecer pautas para la vinculación de la información resultante a la toma de decisiones y a la gestión presupuesto. La concreción de métodos, estándares y cuestiones relativas al uso de los resultados de las evaluaciones se dejan en manos de los diferentes departamentos y agencias, aunque por otro lado se favoreció la creación de unidades de evaluación que incorporaron esas funciones. Se ha tratado de un proceso progresivo de acumulación de experiencia, desarrollado en paralelo a la adopción de técnicas de monitoreo y Performance Measurement propias del NPM, sin que ello haya ido en detrimento de la evaluación. En ese sistema que podríamos calificar de pluralismo centralizado, el Ministerio de Finanzas ha jugado un papel importante de impulso y coordinación, mientras que la Netherlands Audit Office ha asumido la supervisión general del funcionamiento del sistema. En sus últimos informes, pone de manifiesto algunos problemas relativos a la calidad de las evaluaciones, y también márgenes de mejora en cuanto su uso en la toma de decisiones.

18. www.videnet.nl

3. La evaluación en el Reino Unido

El Reino Unido pertenece a la cohorte de países pioneros en el desarrollo de la evaluación de políticas públicas en el entorno internacional (Furubo y Sandahl, 2002). Ello puede atribuirse a la concurrencia de varios factores identificados en apartados previos como determinantes del desarrollo de ésta. En particular, la expansión del *Welfare State* a partir del final de la II Guerra mundial, una larga tradición científica, incluyendo las ciencias sociales, así como la calidad de las instituciones políticas y de la cultura democrática.

No obstante, el desarrollo de la evaluación en Reino Unido se ve frenado a lo largo de la ola de reformas del sector público, de corte gerencial, emprendidas a partir de la década de los 80, en las cuales la evaluación tuvo un papel residual.

No será hasta 1997 cuando vuelva a situarse en la agenda política, evolucionando de forma paralela a las prácticas de *Performance measurement* implantadas durante el período de reformas, sin mantener apenas relación con ellas y teniendo la evaluación todavía un papel más bien subsidiario:

"Evaluation has always been an underdeveloped part of the British policy system and despite definite increases in evaluative activity after 1997, under the rubric of "evidence-based policy," "performance" remains very much the predominant approach." (Talbot, 2010, pág. 1)

"Performance measurement" en el contexto de reforma administrativa

Durante tres décadas, a partir de principios de los 80, los diferentes gobiernos que se han sucedido en el Reino Unido han impulsado diversas políticas orientadas a la implantación de sistemas de *Performance measurement* en el sector público. Este proceso culminó a finales de los 90 con el establecimiento de un sistema prácticamente universal de medición del rendimiento en el gobierno central, autoridades locales, agencias ejecutivas y otros organismos públicos, así como organismos no gubernamentales a cargo de la prestación de servicios públicos.

Ello fue posible gracias a las peculiaridades del sistema político británico, caracterizado hasta hace escasos años por una fuerte concentración de competencias en el gobierno central, sin reconocimiento constitucional de competencias para otros niveles de gobierno. En términos de cobertura poblacional, tras los procesos de *devolution* a Escocia, Gales e Irlanda del Norte, el gobierno central sigue controlando cerca del 80% de la prestación de servicios públicos (Talbot, *op. cit.*).

Hitos importantes en el proceso de reforma son la *Financial Management Initiative* aprobada en 1982, la *Next Steps Initiative* aprobada en 1988 y que llevó a la creación de alrededor de 140 agencias ejecutivas, y las iniciativas *Compulsory Competitive Tendering* (CCT) y *Citizen Charters* o cartas de servicios instauradas en 1991.

La *Financial Management Initiative* de 1982 infiltró una nueva filosofía de gestión descentralizada, incluyendo la gestión del presupuesto, en el gobierno central que se fue reforzando y complementando con las iniciativas posteriores. Desde el principio, ese proceso implicó la adopción de sistemas de medición del rendimiento, incluyendo el establecimiento de objetivos e indicadores, así como métodos para su medición y elaboración de informes periódicos. Más concretamente, el proceso de implantación y expansión del sistema de *Performance measurement* se caracteriza por las etapas siguientes (Wollmann, 2003):

- Unos primeros pasos, a principios de los 80, en algunos entes locales y también en el *National Health Service* (NHS), a través de experiencias aisladas.
- Una fase de desarrollo más sistemático por parte del gobierno central, entre finales de los 80 y principios de los 90, en la que se imponen determinados instrumentos y prácticas. En particular:
 - *Key Performance Indicators* (KPIs), que constituyen la base de los informes anuales que deben elaborar las agencias ejecutivas y presentar a los órganos gubernamentales a los que están adscritos, a partir de 1988.
 - Posteriormente, a partir de principios de los 90, la misma obligación se extiende a otros NDBP (*non-departmental public bodies*) y a los denominados *quangos* (*quasi-autonomous non-governmental organizations*).
 - A partir de 1992, el gobierno central establece la obligación de las autoridades locales de informar anualmente sobre su rendimiento, a partir de un sistema de 200 indicadores estandarizados, organizado y supervisado por la *Audit Commission*. Requerimientos similares de medición del rendimiento se expanden progresivamente a lo largo de esa misma década a una diversidad de servicios públicos como la policía, la educación primaria y secundaria, la educación superior, etc. A mediados de los 90 apenas había organismos a cargo de la prestación de servicios públicos que no estuvieran sometidos a alguna obligación de información periódica sobre sus actividades y resultados (Wollmann, 2003).

A partir de 1997, el *New Labour* promueve un cambio estratégico en el sistema. Por una parte, en el gobierno central se impulsan los *Public Service Agreements* (PSAs) y, que se expanden a través de oleadas sucesivas en 1998, 2000, 2002, 2004 y 2007, en las que se van introduciendo diferentes ajustes. La principal diferencia entre los PSA y el sistema introducido en la etapa anterior lo constituye una mayor orientación estratégica: los PSA expresan objetivos de política pública de los diferentes departamentos del Gobierno británico a medio plazo asociados a asignaciones presupuestarias plurianuales (tres años) -fijadas a través de *Comprehensive Spending Reviews* y de *Spending Reviews*-. En esa misma línea, la medición del grado de progreso debía considerar los *outcomes* o impactos alcanzados, aunque en realidad no ha sido así:

"In 1998 the aim was to cover all departmental objectives and to focus on 'outcomes'. Whilst the former was probably achieved, the latter certainly wasn't." (Talbot, op. cit., pág. 8).

En 2007, el gobierno de Gordon Brown introduce algunos cambios que implican un grado más elevado de complejidad en el sistema: por una parte, los PSA, que en la etapa anterior afectaban normalmente a un único departamento del gobierno, devienen transversales. Al mismo tiempo, se reduce drásticamente su número. Además, cada PSA se complementa con *Delivery Agreements* para precisar responsabilidades y formas de provisión de los servicios públicos implicados. Y, de forma complementaria y a raíz de la drástica reducción en el número de PSA, cada Departamento debe fijar *Departmental Strategic Objectives* (DSOs) en ámbitos no cubiertos por los PSA.

Evaluación

La práctica de la evaluación en el gobierno británico fue relativamente baja durante esta etapa de reformas y hasta 1997, en comparación con otros países de la OCDE (Talbot, *op. cit.*). A partir de 1997, el impulso de la evaluación por parte del gobierno laborista se plasma en diversas iniciativas, en buena parte fomentadas desde unidades especiales adscritas al *Cabinet Office*, a la Oficina del Primer Ministro o a *HM Treasury* (ver Tabla 2). Esas unidades, situadas tradicionalmente fuera de la estructura ordinaria de la Administración, han tenido entre sus funciones la identificación de nuevas formas de hacer las cosas y la difusión de ese conocimiento en los diferentes organismos públicos.

Así como la finalidad de la *Efficiency Unit*, establecida en 1979, fue fundamentalmente mejorar la calidad de la gestión pública en el gobierno, la *Prime Minister's Delivery Unit* (PMDU), creada en 2001, se orientó a mejorar el impacto o *outcome* de las políticas públicas en diferentes ámbitos. La PMDU inició diversas colaboraciones con los departamentos del Gobierno para fijar objetivos, aprender a utilizar datos para entender y tratar de dar solución a los problemas, producir informes regulares, etc.

Por su parte, la *Strategy Unit*, contemporánea de la anterior (creada en 2002), se creó a partir de la fusión de otras dos unidades especiales –la *Forward Strategy Unit* y la *Performance and Innovation Unit*-. Su labor fundamental fue la identificación y análisis de temáticas transversales a los diferentes órganos del Gobierno que implican retos a medio y largo plazo y que requieren de un análisis sofisticado. También contribuyó a la iniciativa de *Spending Reviews* y a la planificación presupuestaria, ayudando a establecer prioridades. En ese marco, llevó a cabo también algunas evaluaciones de impacto de proyectos específicos, en colaboración con ministerios sectoriales.

Tras el cambio de Gobierno en 2010, la *Strategy Unit* fue suprimida y en su lugar se creó el *Behavioural Insights Team*¹⁹, que inicialmente se constituyó como organismo público

19. <http://www.behaviouralinsights.co.uk/>

adscrito al *Cabinet Office* pero actualmente funciona como organismo privado participado por el *Cabinet Office* y por NESTA²⁰, no centrado únicamente en políticas públicas.

Por otra parte, el *Government Social Research Service* (GSRS), a diferencia de las unidades anteriores es una especie de red de investigación pluridisciplinar y evaluación con miembros en los principales organismos del Gobierno. Ha contribuido a impulsar la evaluación y las nuevas tendencias de *evidence-based policy making*. Entre otros productos, ha elaborado el *Magenta Book*²¹, concebido como una guía para evaluadores (sobre perspectivas de evaluación, métodos, comunicación de resultados, etc.), que es revisada y actualizada periódicamente.

Tabla 2. Unidades especiales adscritas al Gobierno del Reino Unido

Primer Ministro	Unidades transversales*	Unidades temáticas	Otras
M. Thatcher (1979-1990)	Central Policy Review Staff Efficiency Unit		Next Steps Unit
John Mayor (1990-1997)	Efficiency Unit	Central Drugs Coordinating Unit	Citizen's Charter Unit
Tony Blair (1997-2007)	Performance and Innovation Unit (1998 – 2002) Prime Minister's Delivery Unit Strategy Unit (inicialmente Forward Strategy Unit)	Social Exclusion Unit (SEU), (trasladada a la Office of the Deputy Prime Minister en 2001) Anti-Drugs Coordination Unit (trasladada al Home Office en 2002) Rough Sleepers Unit (Office of the Deputy Prime Minister) Respect Task Force (in Home Office) Social Exclusion Task Force	Centre for Management and Policy Studies Office for Public Service Reform
Gordon Brown (2007-2010)	Strategy Unit Delivery Unit (trasladada a HM Treasury)		
David Cameron (2010 - actualidad)	Behavioural Insights Team (externalizada en 2013) Implementation Unit	Office for Civil Society Troubled Families Unit	

Fuente: Rutter y Harris (2014). * No se incluye la Policy Unit, creada en 1974 y adscrita al PM desde entonces.

20. NESTA fue creada originalmente por el Gobierno británico en 1998 y actualmente entidad privada sin ánimo de lucro. Su objetivo es promover la innovación basada en evidencias en diversos ámbitos de intervención pública (www.nesta.org.uk)

21. <https://www.gov.uk/government/publications/the-magenta-book>

En 2013, bajo el Gobierno de David Cameron, se puso en marcha la red de *What Works Centres*²², formada actualmente por nueve centros especializados en políticas educativas, de salud pública, servicios penitenciarios, desarrollo local e intervención social temprana²³. En conjunto, las áreas de política pública cubiertas por esta red suponen un volumen total de gasto superior a 270 mil millones de euros. El objetivo de la red es favorecer una toma de decisiones basada en evidencias rigurosas. Con esa finalidad, los centros producen evaluaciones, síntesis de evaluaciones ya existentes, difunden los resultados de forma que sean fácilmente accesibles a los decisores públicos, etc.

Finalmente, un informe reciente de la *National Audit Office* (2013), centrado en evaluaciones de impacto y económicas, pone de manifiesto una práctica desigual:

- Una cobertura difícil de establecer, dada la falta de información sistemática al respecto por parte del Gobierno. Sin embargo, la NAO afirma que entre 2006 y 2012, los departamentos del gobierno publicaron en sus webs un total de 6.000 documentos de análisis, de los cuales 305 eran evaluaciones de impacto, y 70 de ellas incluyeron análisis de coste-efectividad. De esas 70 evaluaciones, 41 afectaron a un total de 12.300 millones de libras de gasto público (16.500 millones de euros).
- Una calidad variable: sólo 14 evaluaciones de impacto, entre una muestra de 34 analizadas por la NAO, habían utilizado metodologías robustas para la definición del contrafactual y presentaban por tanto resultados fiables. Además, las evaluaciones más débiles metodológicamente tendían a mostrar mejores resultados. En general, las evaluaciones relativas a políticas educativas y laborales presentaban una calidad más elevada que las realizadas en otros ámbitos, como políticas urbanas o de promoción empresarial.
- Falta de conexión entre el resultado de evaluaciones previas y análisis *ex ante* o *Impact Assessment* relativos a nuevas fases de las mismas políticas (sólo 40 de 261 *impact assessment* publicados entre 2009 y 2010 hacían referencia a resultados de evaluaciones), y en general entre la evaluación y el proceso de toma de decisiones.
- La forma en que se organiza la función evaluadora varía entre departamentos, sin seguir una pauta clara. Así, mientras es bastante común contar con unidades internas de evaluación, que realizan directamente o encargan evaluaciones expertos, algunos departamentos han establecido organismos con grados variables de independencia para llevar a cabo todas o parte de sus evaluaciones:
 - El *Department for International Development* (DFID) estableció en 2011 la *Independent Commission for Aid Impact* (ICAI), que es un órgano consultivo público, no

22. <https://www.gov.uk/government/publications/what-works-evidence-centres-for-social-policy>

23. El National Institute for Health and Care Excellence (www.nice.org.uk); la Early Intervention Foundation (<http://www.eif.org.uk>); el What Works Centre for Local Economic Growth (<http://whatworksgrowth.org>); What Works Scotland (<http://whatworksscotland.ac.uk>); la Education Endowment Foundation, mencionada anteriormente (<http://educationendowmentfoundation.org.uk>); el College of Policing (<http://www.college.police.uk>); el What Works Centre for Wellbeing (<http://whatworkswellbeing.org>); el Centre for Ageing Better (<http://www.centreforageingbetter.com>), y el Public Policy Institute for Wales (<http://ppi.wales>).

departamental sino adscrito al Parlamento a través del *International Development Committee* de éste.

- El Departamento de Educación ha creado la *Education Endowment Foundation* (EEF)²⁴, que opera como una fundación privada financiada a través de una dotación gubernamental y que es responsable de evaluar intervenciones educativas dirigidas a niños y niñas en situación de riesgo de exclusión social en Inglaterra.
- Por otra parte, según el estudio de la NAO, persisten las barreras tradicionales a la realización de evaluaciones de calidad: problemas de accesibilidad a registros administrativos, discrepancias entre los ritmos y calendarios asociados a los procesos decisionales y a la realización de evaluaciones, falta de demanda de evaluaciones, etc.
- En general, existen todavía márgenes de mejora apreciables en cuanto al uso de los resultados de las evaluaciones en la toma de decisiones

El sistema de evaluación en el Reino Unido

Valoración general

El desarrollo de la evaluación en el Reino Unido parte de unos “activos” similares a los referidos en el caso de Holanda, en particular en lo que respecta a la larga tradición democrática del país, la fortaleza de sus instituciones, su larga tradición científica y en particular el desarrollo de las ciencias sociales. Si bien la realización de evaluaciones rigurosas, y en general el interés de la academia por el estudio de las políticas públicas se remonta, como en Estados Unidos, Escandinavia y Países Bajos al período de entreguerras y persiste con fuerza durante los años dorados de la posguerra, el advenimiento de la crisis del petróleo y el acceso al poder de gobiernos conservadores de corte neoliberal abre un largo paréntesis en la práctica de la evaluación en el gobierno, que no se cerrará hasta el 2000. Durante esos años, se imponen las prácticas de Performance measurement asociadas a la gestión por resultados y al control del gasto, que alcanzarán su apogeo a finales de los 90 habiéndose extendido a todas las agencias del gobierno central, así como a los entes locales. A partir de 2000 vuelve a tomar impulso la evaluación de políticas públicas, de la mano de la Strategy Unit y del Government Social Research Service. Actualmente el Gobierno de David Cameron sigue impulsando la evaluación rigurosa de políticas públicas a través de iniciativas como la red de What Work Centres. Además, existen diversas iniciativas impulsadas por entidades privadas sin ánimo de lucro, aunque con apoyo del Gobierno, muy activas en la promoción del evidence-based policy making, como la Education Endowment Foundation, en Educación, o NESTA, en políticas de innovación y desarrollo y también en innovación local. Sin embargo, a pesar de esa eclosión de iniciativas de promoción de la evaluación, un informe reciente de la National Audit Office revela la existencia de áreas de mejora tanto en la cobertura como en la calidad y en el grado de utilización de las evaluaciones.

24. <http://educationendowmentfoundation.org.uk/>

4. La evaluación en Alemania

En Alemania pueden distinguirse tres olas o etapas en el desarrollo de la evaluación (Speer, *op. cit.*; Derlien, 2002). La primera tuvo lugar entre los años 60 y 70, y se asoció con el desarrollo de las políticas de *Welfare* (Hellstern y Wollmann, citados por Speer, 2012). La segunda oleada se inició a mediados de los 70, en medio de una importante crisis económica y fiscal, y orientada por tanto a introducir criterios para orientar los recortes presupuestarios y para mejorar la eficiencia en el gasto público. La tercera ola vino de la mano del *New Public Management*, designado en Alemania como *New Steering Model*, a principios de los 90, aunque en ese país esta oleada tuvo sólo una conexión débil con la evaluación.

Desde mediados de los 2000 y hasta la actualidad, se está atravesando una cuarta ola o etapa de desarrollo de la evaluación (Speer, 2012). Las nuevas leyes incorporan cada vez más cláusulas sobre la evaluación y sobre la experimentación. Se trata de cláusulas de vigencia limitada, donde la continuidad de leyes y de agencias no está garantizada, sino condicionada a su rendimiento; han sido incorporadas de forma generalizada en los Estados o *Länder*. Al mismo tiempo, los sistemas de evaluación y el flujo de evaluaciones de programas se han ido extendiendo a ese nivel, especialmente en lo que respecta a las evaluaciones de impacto.

Las instituciones

Alemania es una república federal compuesta por 16 estados o *Länder*. Hay equilibrios de poder y distribución de responsabilidades entre el nivel federal y los estados que varían en función del ámbito de política pública, lo que influye también en cómo se desarrolla la evaluación. Es el caso de las políticas de empleo, que se diseñan en el nivel nacional, y las políticas educativas, que se gestionan de forma más descentralizada, con implicación tanto de los *Länder* como de los gobiernos locales y donde, por tanto, se da una gran diversidad, también en lo que respecta a la práctica de la evaluación.

Por otra parte, los ministerios disfrutaban de un alto grado de autonomía, lo que influye también en la falta de un modelo de institucionalización claro de la evaluación:

"Co-operation and consensus building are also key features of the way in which the federal executive works. The principle of ministerial autonomy means that the chancellery acts more as a co-ordinator than a driver of policy and law making. Centrifugal forces need to be kept in check and the system raises a significant challenge for the centralisation of reform, the establishment of clear reform leadership in the executive centre of the federal government, and the development of a collective, whole-of-government approach to reform". (OECD, 2010, pág. 60).

En Alemania, la evaluación se ha incorporado de una forma más procedimental que institucional. La influencia nacional se hace notar fundamentalmente a través de las leyes y en el ámbito de discrecionalidad de los diferentes Ministerios. No existe un marco institucional global de la evaluación a nivel estatal. El Tribunal de Cuentas (*Bundesrechnungshof*) tiene una influencia limitada en la evaluación, ya que ésta se desarrolla fundamentalmente en el marco de las políticas sectoriales, donde diferentes reformas y cambios en las políticas han traído de la mano una mayor demanda de evaluaciones y los fondos necesarios para financiarlas. Por su parte, el Parlamento requiere de forma creciente información sobre evaluaciones de políticas públicas, aunque de momento las interpretaciones al respecto se centran más bien en cuestiones procedimentales (qué evaluaciones se encargan, a qué entidad, con qué calendario, etc.) que en los resultados de los análisis (Jacob, Speer y Furubo, 2015).

A pesar de la falta de un sistema de institucionalización formal de la evaluación, es indudable la existencia de una cultura de evaluación madura, generada y a la vez compartida por una comunidad de conocimiento y práctica de la evaluación altamente desarrollada. Algunos indicadores al respecto son la existencia de DeGEval, una asociación alemana y austríaca de evaluación muy activa²⁵. También, una revista específica sobre evaluación, de periodicidad semestral (*Zeitschrift für Evaluation*), así como la existencia de unos estándares de evaluación aprobados por DeGEval ya en 2001.

Las prácticas

Las evaluaciones de políticas públicas se llevan a cabo de forma regular en diferentes ámbitos de intervención pública. En particular, en el ámbito educativo y en el de las políticas de empleo, así como en el de la cooperación internacional para el desarrollo (Speer, *op. cit.*). Sin embargo, en la práctica de la evaluación hay una separación más profunda que en otros países europeos entre el sector de investigación académica y la Administración. Así, por ejemplo, la evaluación de las políticas de empleo se lleva a cabo por institutos de investigación, aunque en muchos casos adscritos a la Administración.

Es el caso del *Institute for Employment Research* (IAB)²⁶, un instituto de investigación adscrito a la Agencia Federal de Empleo (*Bundesagentur für Arbeit* – BA). Se fundó en 1967 y la evaluación de políticas de empleo está entre sus objetivos principales. Los objetivos de las evaluaciones que lleva a cabo la BA, a través del IAB, incluyen: la mejora de las perspectivas de empleo y de la empleabilidad de los beneficiarios; los costes y beneficios de los programas; la medida de los efectos macroeconómicos netos de los programas, y el impacto regional de las políticas de empleo.

25. <http://www.degeval.de/home/>

26. <http://www.iab.de/en/iab-aktuell.aspx>

Aunque adscrito a la BA, el IAB es independiente y es libre de publicar las evaluaciones y de hacerlo en el momento que juzgue más adecuado. Su naturaleza de instituto de investigación preserva su autonomía al respecto, así como la aplicación de métodos y estándares rigurosos de evaluación. El Ministerio Federal de Empleo tuvo durante menos de un año una unidad de evaluación, que se suprimió (Speer, *op. cit.*). Aunque el IAB lleva a cabo la mayor parte de las evaluaciones de políticas de empleo, hay también otros actores como institutos públicos de investigación u organizaciones o grupos de investigación académica, que también llevan a cabo evaluaciones en este sector.

En lo que respecta a las políticas educativas, existe un elevado número de instituciones involucradas en su evaluación. En el nivel federal, por ejemplo, el *Leibniz Institute for Educational Research and Educational Information* (DIPF, *Deutsches Institut für Internationale Pädagogische Forschung*) es el actor principal. En el nivel de los *Länder*, diversos ministerios financian institutos de investigación, a menudo adscritos a universidades locales. Por otra parte, se han creado recientemente en algunos casos agencias de inspección educativa que dependen de los ministerios de los *Länder*. También se ha puesto en marcha un Instituto para el Desarrollo de la Calidad de la Educación (IQB), afiliado a la Universidad Humboldt de Berlín y que se ocupa del desarrollo de estándares educativos.

El Ministerio estatal de Educación ha aumentado los fondos para investigación educativa, así como para la evaluación de políticas educativas. Las escuelas disponen de presupuesto para actividades de evaluación, que aunque de cuantía limitada, pueden administrar de forma flexible, determinando ellos mismos las evaluaciones que estiman necesarias. Hay un mercado bastante activo de empresas de consultoría que cubren este tipo de evaluaciones (Speer, *op. cit.*).

La repercusión pública de los resultados de las evaluaciones es mayor en el caso de la educación que en el de las políticas de empleo, donde el debate queda restringido a grupos de expertos y *practitioners*, sin apenas repercusión en los medios de comunicación.

Por otra parte, las comunidades de evaluación son diferentes en ambos casos, y se encuentran claramente separadas: psicólogos, directores de escuela y maestros, en el caso de la Educación; investigadores académicos, y especialmente economistas, en el caso de las políticas de empleo. En palabras de Speer:

"In Germany, where the institutionalization of evaluation on the national level is relatively lower, the evaluation cultures are very unique to each sector... Although an active national evaluation society exists in Germany, the communities carrying out the evaluations remain fragmented and there is a lack of interdisciplinary discourse on the subject". (Speer, *op. cit.*, pos. 1787)

Evaluación en los gobiernos locales

Tampoco en el caso de Alemania hay estudios relevantes sobre la institucionalización y la práctica de la evaluación en el nivel local. Si se describe, en cambio, la adopción y evolución de sistemas de monitoreo y de *Performance management* y *Performance monitoring*, introducidas de la mano del *New Steering Model* a lo largo de la década de los 90.

En Alemania, la situación de estas prácticas a nivel municipal muestra una curiosa mezcla de los elementos que se encuentran en ese mismo nivel en Suecia, Francia y Reino Unido. Por una parte, como en el caso de Suecia y Francia, su surgimiento sigue una dinámica de abajo arriba, aunque consiguiendo un amplio seguimiento: según Kuhlmann, a lo largo de la década de los 90, alrededor del 82% de los gobiernos locales alemanes emprendieron reformas de ese tipo. Como en el caso de Francia, en la mayor parte de los casos se adoptaron sistemas de indicadores centrados sobre todo en inputs, actividades y resultados.

Se pusieron también en marcha dispositivos de comparación intermunicipal (de costes y calidad de los servicios, entre otros aspectos) que involucraron el 60% de los municipios alemanes. El sistema resultante se asemeja al británico en su complejidad y parece excesivamente burocrático en algunos casos:

"There is a precarious tendency of adding to the already existing bureaucracy of legal rules and procedures a new one that consists of indicators and 'products'. This 'bureaucratic out-growth' of performance indicators in Germany is illustrated by the management chart for the citizen services authority (Bürgerservice). The KGSt proposes 66 performance indicators with detailed measurement instructions." (Kuhlmann, op. cit. pág. 339)

Por otra parte, las autoridades locales alemanas son libres de publicar o no la información recogida a través de esos dispositivos. Las comparaciones intermunicipales no acostumbran a publicarse y, en general, la información producida apenas se utiliza:

"In practice, performance management instruments are barely used for political or administrative decision-making. In most cases, the information about the products and performance is neither applied for budget approval nor for the reorganization of administrative procedures or inter-municipal comparison. Most strikingly, 14 percent of local authorities that deal with performance indicators do not use this information at all." (Kuhlmann, 2010, pág. 339-340)

El sistema de evaluación en Alemania

Valoración general

Alemania sigue una pauta muy similar de evolución de la evaluación que los tres países vistos anteriormente hasta mediados de 2000: presencia temprana, impulso en los años dorados de la posguerra y cambio de rumbo en los 80 y 90 de la mano del New Steering Model, hacia fórmulas gerenciales de monitoreo y medición del rendimiento. Desde 2000 se abre paso una etapa de institucionalización más procedimental que formal, aunque incluye la regulación a través de cláusulas legislativas que imponen obligaciones de evaluación en determinados ámbitos. La estructura federal del estado y el elevado grado de autonomía que disfrutaban los ministerios ha favorecido una implantación diversificada de la evaluación, con mayor presencia en determinados sectores de política pública, como es el caso de la Educación y de las políticas de empleo. En Alemania la evaluación se lleva a cabo normalmente de forma externa a la Administración, en institutos de investigación bien vinculados orgánicamente a los organismos responsables de la política, o bien independientes. En cualquier caso, la evaluación se realiza siguiendo elevados estándares metodológicos y con plena autonomía para ejecutar los estudios y también para difundirlos. Existe, sin embargo, una elevada compartimentación entre disciplinas académicas, así como entre academia y administración. Por otra parte, el Tribunal de Cuentas Alemán tiene una influencia limitada en este ámbito, a diferencia del Reino Unido y los Países Bajos. El uso de las evaluaciones a efectos de mejora de las políticas parece elevado en el nivel federal.

5. La evaluación en Francia

El proceso de institucionalización de la evaluación en Francia comienza en los años 80, de la mano de iniciativas de reforma de la administración impulsadas por sucesivos gobiernos socialistas, bajo la presidencia de François Mitterrand (1981-1995). Después del fracaso de la operación *Rationalisation des Choix Budgétaires* (RCB)²⁷, la *Direction de la Prévision* del Ministerio de Economía, Finanzas y Presupuesto trata de transformar una red interministerial constituida a raíz de la experiencia RCB en una red de evaluación, para aprovechar de ese modo la experiencia acumulada en análisis de políticas públicas.

Con la finalidad de lanzar la iniciativa y de sensibilización, en Diciembre de 1983 se celebra en París un coloquio sobre evaluación de políticas públicas, organizado conjuntamente por el Ministerio de Economía, Finanzas y Presupuesto y la *École des Hautes Études Commerciales*. A partir de ese momento, la evaluación será objeto de consideración en la Administración pública francesa, aunque sin demasiado entusiasmo²⁸.

27. Adaptación francesa del *Planning, Programming & Budgeting System* (PPBS), introducido en el Gobierno Federal Norteamericano bajo la Presidencia de Jimmy Carter en 1979.

28. "Le Commissariat général du Plan n'exprime qu'une tiède bienveillance. Plus incisive, la Direction du Budget fait savoir à la DP [Direction de la prévision] qu'elle n'est pas favorable à une relance du dispositif RCB, même sur une base

De forma posterior a ese coloquio, la cuestión de la evaluación es retomada por dos ministros: Jacques Delors (Economía y Finanzas) y Michel Rocard (Plan), quienes encargan un informe a Michel Deleau sobre la necesidad de institucionalizar su práctica. Un grupo de trabajo formado por expertos externos participa en esa reflexión, que finaliza con la presentación del informe final *Évaluer les politiques publiques: méthodes, déontologies, organisation: rapport*, en 1986. Poco después de la presentación del informe, se produce un cambio de gobierno, con Chirac como Primer Ministro y Mitterrand como Presidente de la República (primer período de cohabitación). El nuevo Ministro de Finanzas, Édouard Balladur, se abstendrá de aplicar y dar seguimiento al informe.

En 1988, con Michel Rocard como Primer Ministro, se vincula la modernización de los servicios públicos a la evaluación, que se concibe como la contrapartida indispensable de la autonomía. Con esa finalidad, se encarga un nuevo informe externo que, bajo el título *L'Évaluation des politiques publiques et des actions publiques*²⁹, servirá de base para la circular de 23 de Febrero de 1989 *Renouveau du service public*. Este último informe enfatiza una aproximación de control democrático, definiendo la evaluación como un juicio realizado por representantes elegidos democráticamente, sobre el valor de las políticas públicas (Jacob, *op. cit.*; Fontaine y Monnier, 2002). Se pone el acento, por tanto, en la necesidad de que multiplicar los focos de evaluación para reforzar el equilibrio de poderes, especialmente en relación al ejecutivo, para evitar que éste monopolice el conocimiento experto sobre las intervenciones públicas.

El dispositivo de evaluación interministerial que deriva de la circular mencionada constituye el primer marco institucional de la evaluación en Francia. Se compone de cuatro instancias: el *Commissariat général du Plan*³⁰, el *Fonds national du développement de l'évaluation*, el *Comité interministériel de l'évaluation* (CIME) y el *Conseil scientifique de l'évaluation*. Las evaluaciones promovidas en ese marco, y en particular por el *Conseil scientifique de l'évaluation* adoptan un carácter pluralista, siendo encomendadas a comités de evaluación compuestos por cargos electos, gestores públicos y grupos de beneficiarios, que deben formular un juicio de valor sobre las políticas analizadas (Fontaine y Monnier, 2002).

Las evaluaciones realizadas son numerosas, pero sus recomendaciones carecen de seguimiento (Jacob, *op. cit.*). Además, a media que transcurre el tiempo van elevándose cada vez más críticas contra un mecanismo considerado demasiado rígido tanto por

évaluative. Pour elle, l'échec de la RCB montre que les politiques publique se gèrent à coup de rapports de force où la raison intervient peu." (Spenlehauer, citado por Jacob, *op. cit.* pág. 52).

29. Viveret, P., *L'évaluation des politiques publiques et des actions publiques*; seguido de *Propositions en vue de l'évaluation du revenu minimum d'insertion : rapports*, Paris, la Documentation française, 1989.

30. Institución pública creada en 1946, estuvo a cargo de la planificación económica de Francia, a través de la definición de planes quinquenales, hasta 2006. Después de esa fecha se transformó en el Centre d'Analyse Stratégique (CAS).

altos funcionarios como por investigadores. En general, se considera que la regulación de los trabajos es de un excesivo rigor y complejidad científicos, cuando apenas existen funcionarios preparados para utilizarla³¹.

El desinterés progresivo de los decisores públicos por los trabajos del Comité Intermi-nisterial de Evaluación (CIME) se pone de manifiesto por la ausencia de convocatorias desde Marzo de 1993, la ausencia de encargos de evaluación y la no renovación de sus miembros a medida que van agotando sus mandatos. Finalmente, a finales de 1998 se modifica el dispositivo de evaluación en el poder ejecutivo, haciéndolo reposar sobre el *Conseil national de l'évaluation*, y el *Commissariat général du Plan*, como responsable de la evaluación de las políticas públicas estatales y locales. El *Conseil national de l'évaluation* se compone de catorce miembros, designados por un período de tres años y que pueden ser representantes de organismos nacionales, de organismos locales, o bien expertos en evaluación. Una de las funciones principales del *Conseil* era proponer un programa de evaluaciones anual y proveer asistencia técnica y formación para su implementación por parte de los organismos públicos competentes en las políticas a evaluar. También debía supervisar la calidad de las evaluaciones y ocuparse de su difusión pública. Sin embargo, el nuevo dispositivo interdepartamental recoge también escaso éxito (Jacob, *op. cit.*).

A pesar de los fracasos sucesivos en los mecanismos interdepartamentales de evaluación, ésta se va desarrollando en determinados ámbitos de intervención, como es el caso de la Educación. Así, desde 1987, se producen iniciativas sucesivas de institucionalización de la evaluación en el ámbito educativo como lo muestra la creación de la *Direction de l'évaluation et de la prospective* en el seno del *Ministère de l'Éducation*. A principios de 2000 el nuevo ministro de Educación, Jack Lang, inaugura el *Haut Conseil pour l'évaluation de l'école*. También se crea, a finales de los 80, el Comité Nacional para la Evaluación de las Políticas Urbanas, así como diversos órganos de evaluación en los ámbitos de la investigación científica y de la formación profesional.

En el sector de las políticas sociales, tiene especial importancia la introducción de una cláusula de evaluación en la Ley sobre la Renta Mínima de Inserción, adoptada por la Asamblea francesa en Diciembre de 1988. En concreto, el texto contiene una disposición que subordina, al término de un período de tres años, la continuación del programa a los resultados de su evaluación. Sin embargo, al finalizar el período establecido en la norma, ésta es modificada y se abandona la exigencia de evaluación. No obstante, una versión posterior, la *Loi de la Revenu de Solidarité Active* aprobada a finales de 2008 servirá de base para la puesta en marcha de una iniciativa innovadora en materia de evaluación, que llega hasta la actualidad: el *Fonds d'Expérimentation pour la Jeunesse*³².

31. "Quand on pose l'évaluation des politiques publiques dans le sacré, ce n'est pas étonnant qu'on n'en fasse pas." (Audebrand, É., citado por Jacob, *op. cit.*, pág. 79).

32. <http://www.experimentation.jeunes.gouv.fr/>

En concreto, la norma establece que se dotará al Fondo de contribuciones del estado y otras entidades que se asocien para definir, financiar y pilotar uno o más programas experimentales orientados a mejorar la inserción social y profesional de los jóvenes de 16 a 25 años. El Fondo financia los programas, y también evaluaciones independientes de impacto que deben asociarse obligatoriamente a los mismos. A finales de 2010 se amplía su finalidad y también la población diana, y se establece que el Fondo se destinará a favorecer el éxito escolar de los alumnos y mejorar la inserción social y profesional de los jóvenes de menos de 25 años (Gurgand y Valdenaire, 2012). Se trata de una iniciativa interesante, que se enmarca en nuevas tendencias de innovación social apoyadas en evaluaciones rigurosas de impacto. Pero su interés radica también en la financiación y en la gestión mixta, público-privada del Fondo (Lázaro, 2012b).

En lo que se refiere al papel del Parlamento francés, en un primer momento este pone en funcionamiento una instancia común a la Asamblea Nacional y al Senado: se propone, en 1995, a través de una proposición de ley, la creación de una Oficina parlamentaria de evaluación de las políticas públicas (OPEPP), asignándole evaluaciones de tipo tanto prospectivo como retrospectivo sobre la adecuación entre los medios jurídicos, administrativos o financieros consagrados a una política pública y los resultados observados o previstos. La Oficina se componía de dos delegaciones, una de la Asamblea Nacional y otra del Senado, y recibió diversas críticas porque sus trabajos se solapaban con los de las comisiones parlamentarias, y también porque el "expertise" interno estaba aislado del proceso de toma de decisiones. Finalmente, esta primera experiencia del Parlamento, de dotarse de medios propios de evaluación de políticas públicas, fracasa y la Oficina es suprimida en 2001.

Posteriormente, cada Cámara establece una instancia propia de evaluación. El Senado, que desea dar continuidad a los trabajos de la OPEPP, instituye un *Comité d'évaluation des politiques publiques*. En una primera fase, se compone de 16 miembros del Senado que habían participado en los trabajos de la OPEPP. Posteriormente, se adscribe a la Comisión de Finanzas. Los trabajos de evaluación que impulsa la Comisión son llevados a cabo por expertos externos y dan lugar a comparecencias de los responsables administrativos a cargo de los programas analizados.

Por su parte, la Asamblea Nacional opta por otro mecanismo, la *Mission d'évaluation et de contrôle* (MEC) que se crea a raíz de un grupo de trabajo sobre la efectividad del gasto público y el control parlamentario. La MEC tiene por objetivo ejercer las prerrogativas reconocidas al Parlamento en materia de control de la utilización de los créditos incluidos en la Ley de Presupuestos. Se crea en Febrero de 1999, y para agilizar esa creación no se pone en marcha como una comisión permanente, sino que debe renovarse cada año en el seno de la Comisión de Finanzas. Ese procedimiento presenta la ventaja de obligar a los miembros de la Asamblea a expresar cada año su voluntad de proseguir con los trabajos, lo cual es interesante a

efectos de compromiso explícito con la evaluación. Por otra parte, más que encargar nuevas evaluaciones, la MEC realiza análisis de evaluaciones ya existentes y difunde sus contenidos, actuando de caja de resonancia, aunque con un éxito limitado³³.

Actualmente, la Asamblea Nacional cuenta con tres dispositivos de evaluación:

- La *Mission d'évaluation et de contrôle* (MEC), que continua desde 1999 y que cumple un papel inspirado en la *National Audit Office* británica, de control de la efectividad y eficiencia del gasto público, pudiendo llevar a cabo evaluaciones en profundidad de políticas sectoriales.
- La *Mission d'évaluation et de contrôle des lois de financement de la sécurité sociale* (MECSS), creada a finales de 2004 para supervisar el uso dado a los fondos de la Seguridad Social, a través de auditorías, estudios, etc.
- El *Comité d'évaluation et de contrôle* (CEC), creado en Mayo de 2009 y encargado expresamente, entre otras funciones, de la evaluación de políticas públicas en ámbitos que afectan a más de una comisión parlamentaria.

El Senado no parece contar con ningún órgano específico de evaluación, aunque sí dispone de un programa específico de evaluación y prospectiva³⁴ estructurado en diferentes temas, encargados a diferentes comisiones.

En cuanto al papel de la *Cour des comptes*, en Francia el rol que juega este organismo con respecto a la evaluación es ambiguo. A diferencia de una implicación muy activa en el caso de Suiza y Países Bajos, en Francia presenta, por una parte, un discurso muy favorable a las actividades de evaluación, pero en la práctica es bastante más prudente (Jacob, *op. cit*), evitando entrar a valorar la oportunidad de las decisiones tomadas por el Ejecutivo y el Legislativo y centrándose en cambio en analizar la coherencia entre éstas y las actuaciones realizadas. Sin embargo, a parir de la reforma constitucional de 2008³⁵, la *Cour des comptes* colabora directamente con la *Mission d'évaluation et de contrôle* (MEC) de la Asamblea nacional.

33. "Les suites accordées aux travaux de la MEC sont très limitées et il arrive même que le Gouvernement prenne le contre-pied des suggestions qu'elle formule." (Jacob, *op. cit*, pág. 76).

34. Programme de contrôle, d'information, d'évaluation et de prospective du Sénat. <http://www.senat.fr/contrôle/2014/organismes.html>

35. En Julio de 2007 se pone en marcha un Comité de réflexion et de proposition sur la modernisation et le rééquilibrage des institutions, presidido por Édouard Balladur. Los trabajos del Comité darán pie, entre otros, a un proyecto de ley orgánica que se aprobará por la Asamblea Nacional en Julio de 2008 y que revaloriza el papel del Parlamento en el control del Ejecutivo. En particular, la Ley (*loi constitutionnelle no 2008-724 du 23 juillet 2008 de modernisation des institutions de la Ve République*) modifica el Título V de la Constitución, introduciendo un nuevo artículo 47.2 que establece que "La Cour des comptes assiste le Parlement dans le contrôle de l'action du gouvernement. Elle assiste le Parlement et le gouvernement dans le contrôle de l'exécution des lois de finances et de l'application des lois de financement de la sécurité sociale ainsi que dans l'évaluation des politiques publiques. Elle contribue à l'information des citoyens via des rapports publics".

La evaluación en los gobiernos regionales y locales

En Marzo de 1982, se promulga la Ley de derechos y libertades de los municipios, departamentos y regiones, que establece la descentralización de competencias del gobierno central en favor de esos órganos de la administración local. La nueva ley establece, entre otras disposiciones, la realización de contratos programa entre el Estado y las regiones, lo cual, junto con las obligaciones derivadas de la gestión de fondos estructurales europeos, favorece el florecimiento de la evaluación en las administraciones regionales y la creación de unidades organizativas específicas con esa función. Sin embargo, igual que en el caso de los municipios, las formas de institucionalización y las prácticas de evaluación son heterogéneas. Por ejemplo, la Región de Rhône-Alpes estableció en 1991 un Comité Regional de Evaluación, como foro de deliberación sobre las actuaciones a evaluar así como sobre sus resultados, y un comité científico para garantizar la corrección metodológica. En Bretaña, en cambio, se han creado órganos similares pero con participación estatal. A lo largo de la década de los 90 los dispositivos de evaluación se fueron extendiendo por diferentes regiones (Midi-Pyrénées en 1992, Pays-de-la-Loire en 1993, Lorraine en 1994, etc.).

En lo que respecta a los municipios, tampoco se ha puesto en marcha ningún dispositivo específico y homogéneo desde la Administración central. Sin embargo, éstos se han mostrado especialmente activos (Kuhlmann, *op. cit.*). Hay que señalar que Francia presenta una elevada fragmentación de su administración local, con 36.000 municipios, el 75% de los cuales tiene menos de 1.000 habitantes. Por ese motivo, el avance de la evaluación se ha producido de forma irregular, “a dos velocidades”: más rápido en las grandes ciudades, los órganos de cooperación intermunicipal y los departamentos, y más lento o nulo en municipios medianos y pequeños. En cualquier caso, se ha tratado de un proceso voluntario, “de abajo arriba”, heterogéneo y centrado especialmente en sistemas de monitoreo, más que en evaluación, bien adaptados al contexto local y a los servicios concretos a los que se aplica, aunque centrados en indicadores fáciles de medir, de *input*, actividad, y resultados, sin apenas presencia de indicadores de impacto. El arraigo local de esos sistemas favorece su utilización interna, pero complica las comparaciones intermunicipales. Y, en general hay resistencia a la difusión pública de resultados (Kuhlmann, *op. cit.*).

La profesionalización de la evaluación

La *Société française de l'évaluation* se constituye formalmente en un primer congreso celebrado en Marsella en Junio de 1999. La SFE juega un papel pedagógico y de promoción de la evaluación. Organiza un coloquio anual que reúne personal de la Administración, de la academia y del sector de consultoría privada. La SFE ha instituido diversos grupos de trabajo sobre aspectos diversos de la evaluación (métodos, práctica, capacidad de evaluación, etc.), y también secciones regionales.

El sistema de evaluación en Francia

Valoración general

La evaluación se desarrolla en Francia a partir de los años 80, en intentos sucesivos de institucionalización formal tanto en el poder ejecutivo como en el legislativo que, tras diversos ensayos y fracasos, han dado lugar a una práctica pluralista y descentralizada. La evaluación ocupa un lugar destacado en el discurso político vinculado a la reforma de la Administración de los años 80, aunque los intentos de institucionalización no llegan a cuajar en esos años. El primer dispositivo de evaluación se pone en marcha en el Ejecutivo en 1990, tiene carácter interministerial y reposa sobre cuatro pilares: el Commissariat général du Plan, el Fonds national du développement de l'évaluation, el Comité interministériel de l'évaluation (CIME) y el Conseil scientifique de l'évaluation. Se realizan numerosas evaluaciones pero sus recomendaciones no se utilizan y el dispositivo es considerado demasiado rígido. Finalmente, a finales de 1998 se simplifica, haciéndolo reposar únicamente sobre el Conseil national de l'évaluation, y el Commissariat général du Plan. El Conseil national de l'évaluation integraba representantes de organismos nacionales, de organismos locales y expertos en evaluación. Debía impulsar un programa de evaluaciones anual y proveer asistencia técnica y formación para su implementación por parte de los organismos públicos. También debía supervisar la calidad de las evaluaciones y ocuparse de su difusión pública. Sin embargo, el nuevo dispositivo interdepartamental recoge también escaso éxito. A pesar de los fracasos sucesivos en los mecanismos interdepartamentales de evaluación, ésta se va desarrollando en determinados ámbitos de intervención, como la Educación, las políticas sociales –a partir de una cláusula de evaluación en la Ley de la renta mínima de inserción– y las políticas urbanas, incluyendo evaluaciones rigurosas de impacto. Por otra parte, en el Legislativo la Asamblea Nacional cuenta con diferentes instancias de evaluación, aunque su repercusión a efectos de control de ejecutivo parece limitada, como también lo es el papel de la Cour des Comptes.

6. La evaluación en Bélgica

En el caso de Bélgica, tal como afirman Varone, Jacob y De Winter (2005), la evaluación de políticas públicas tiene todavía una presencia débil, tanto en términos institucionales como en lo que se refiere a la existencia de una cultura favorable. Contribuyen a esa situación las peculiaridades del sistema político, y en particular a la debilidad del sistema parlamentario y la centralidad de los partidos políticos en la definición de la agenda política. Ambos factores, entre otros, interfieren en el funcionamiento de los mecanismos de rendición de cuentas y en la institucionalización de la evaluación.

La fragmentación de la sociedad belga por motivos confesionales, socioeconómicos y culturales se traduce en un sistema político multipartidista y en la formación habitual de gobiernos de coalición. Se trata de un ejemplo de democracia consociativa, con profundas divisiones entre sub-culturas diversas. Las coaliciones de gobierno preservan

su frágil equilibrio minimizando la interacción con actores políticos externos a la coalición. El proceso de elaboración de las políticas es “capturado” por los gabinetes ministeriales, constituidos por personas de confianza de los partidos (Varone, Jacob y de Winter, 2005).

Este contexto es desfavorable a la evaluación en dos sentidos. En primer lugar, es refractario a la introducción de factores técnicos que puedan contradecir posiciones de partido y poner en riesgo la estabilidad de la coalición en el gobierno. Por otra parte, el control de la iniciativa legislativa por parte del Gobierno minimiza la posibilidad de introducir disposiciones que favorezcan una evaluación independiente. En segundo lugar, aunque la implementación de las políticas recae en la administración, ésta no tiene ni competencias ni recursos para impulsar por sí misma su evaluación, y en caso de que ésta se lleve a cabo, las posibilidades de utilización y de difusión pública son remotas. En este entorno, el impulso de la evaluación parece depender directamente de la voluntad de los partidos en el gobierno. No obstante, su impulso se encuentra en gran parte ausente de los programas y de las campañas electorales (Varone, Jacob y de Winter, 2005).

Formalmente sí se han dado algunos pasos, si bien en el caso de Bélgica las reformas administrativas inspiradas en el *New Public Management* han olvidado en gran medida la evaluación (Jacob, op. cit.), de modo que, a pesar de algunas iniciativas aisladas, la evaluación de políticas públicas continúa siendo algo raro, escasamente sistemático y disperso en el gobierno y en la Administración belga, sobre todo debido a una toma de conciencia insuficiente sobre su utilidad.

En un primer momento, la evaluación se emplea para buscar el consenso en determinados problemas políticos (Jacob, op. cit.). Un ejemplo emblemático lo proporciona la tramitación del proyecto de ley de interrupción del embarazo en 1990, en el cual la evaluación asume un papel de “garantía” desde la perspectiva de la oposición, como instrumento de control de la implementación de la ley y de alerta sobre posibles desviaciones. Así, la ley en cuestión, promulgada en Agosto de 1990, establece la creación de una Comisión de Evaluación.

Esa misma perspectiva se aplicará en 1999, en la ley relativa a la cooperación internacional belga, promulgada en Mayo de ese año. Esa norma establece la obligatoriedad de presentar informes anuales de evaluación al Parlamento, y se crea un servicio de evaluación en el seno de la Dirección General de Cooperación Internacional. Otro ejemplo es la creación, en 2002, de la Comisión federal de control y de evaluación de la ley que regula la despenalización de la eutanasia en determinados supuestos. Tal como dice Jacob, en esos casos

“Le recours à l'évaluation apparaît comme un instrument “d'apaisement” au cours de débat éthiques.” (Jacob, op. cit., pág. 71)

En 1999, a partir de la formación de un Gobierno “arcoíris” liberal, socialista y ecologista, una declaración gubernamental pone el acento en la evaluación de la legislación:

*“Le Gouvernement, avec le Parlement et en collaboration avec le bureau de coordination du Conseil d’État, procédera à une évaluation de la législation. Celle-ci est un moyen d’améliorer l’efficacité, l’adhésion citoyenne et la qualité juridique de la législation. Cette évaluation de la législation permettra de juger leur fonctionnement et leur efficacité, de sorte qu’elles puissent être revues et, au besoin, abrogées.”*³⁶

Sin embargo, la institucionalización de la evaluación no fue finalmente incluida en la agenda del Gobierno Verhofstadt en esa primera etapa, entre 1999 y 2003.

En cuanto a instancias organizativas responsables de la evaluación, destacan dos. Por una parte, en el Ejecutivo, el *Bureau fédéral du Plan*, que se encarga, después de una reforma legislativa de sus funciones ocurrida en 1994, de analizar y de prever la evolución socioeconómica así como los factores que la determinan, y de evaluar las consecuencias de las decisiones en materia de política económica, con vistas a mejorar la racionalidad, la eficacia y la transparencia de dichas políticas³⁷. La Oficina se sitúa bajo la autoridad del Primer Ministro y del Ministro de Economía, y sus líneas de actuación son refrendadas por el Consejo de Ministros. Es importante destacar que, a diferencia de otros países como los Países Bajos, Francia o Suiza, en Bélgica no hay ninguna tentativa de crear dispositivos de institucionalización transversal de la evaluación en el Ejecutivo (comités interministeriales o similares).

En 2001, la *Association universitaire de recherche sur l’action publique* de la *Université catholique de Louvain* realiza un estudio sobre la práctica de la evaluación en la Administración belga. El estudio muestra que una mayoría de los organismos públicos recurre a la evaluación como un instrumento de gestión. Las evaluaciones son realizadas a iniciativa de los órganos administrativos, a menudo en ausencia de un mandato político expreso, y son ejecutadas y utilizadas de forma interna. Generalmente, ni los decisores políticos ni los beneficiarios de las políticas son implicados en las evaluaciones.

Por otra parte, en 1998 el Tribunal de Cuentas (*Cour des Comptes*), adscrito al Parlamento, amplió sus competencias a la comprobación de la eficacia y la eficiencia en el uso de los recursos públicos³⁸. Las evaluaciones realizadas desde entonces han tenido, sin embargo, escasa resonancia. La insuficiente oferta de evaluadores cualificados, algunas limitaciones en términos de calidad técnica y acceso a información, así como la

36. « La voie vers le XXI^e siècle ». *Déclaration gouvernementale prononcée devant le Parlement le 14 juillet 1999*, Bruxelles, SFI, 1999, P.22. Citada por Jacob, op. cit., pág. 51.

37. Ley de 21 de Diciembre de 1994, *Moniteur belge*, 23 de Diciembre de 1994.

38. El rol de la *Cour des comptes* belga con respecto a la evaluación se define en la ley de 10 de marzo de 1998, que modifica regulaciones precedentes y en la que se establece que el Tribunal debe controlar el buen uso de los fondos públicos.

ausencia de implicación de partes interesadas en las evaluaciones y de prácticas eficaces de difusión de resultados –junto con los factores institucionales vistos anteriormente– pueden haber contribuido a ese resultado.

En lo que respecta al Parlamento, en general, en Bélgica, como en Holanda, el legislador no tiene un rol protagonista en el desarrollo de la evaluación (Jacob, *op. cit.*), más allá de aprobar cláusulas legislativas específicas sobre evaluación, como hemos visto más arriba. No crea, sin embargo, ninguna instancia interna *ad hoc* dedicada a la evaluación.

Por último, aunque algunos autores (Varone, Jacob, de Winter, 2005) sugieren que la estructura federal del estado puede actuar como barrera adicional para la evaluación –debido a una mayor fragmentación y complejidad en los procesos de elaboración e implementación de políticas, a conflictos de competencias que dificultan la coordinación y la transparencia, etc.–, reconocen que puede jugar también a favor al ofrecer un escenario idóneo como “laboratorio” para la experimentación territorial y la evaluación comparada de programas. Estiman, además, que el ejemplo de la región de Valonia –con prácticas de evaluación más avanzadas, desarrolladas al calor de las obligaciones asociadas a la recepción de fondos estructurales europeos– puede actuar como acicate para el resto del país.

La implantación de la evaluación en la región es progresivo, haciendo énfasis en introducir un cultura favorable, que se apoya en una *Société Wallonne d'évaluation* (SWEP), creada en 1999, y en un *Institut wallon de l'évaluation, de la prospective et de la statistique*³⁹(IWEPS) puesto en marcha en 2004. La finalidad del IWEPS es constituirse como centro de investigación independiente, producir conocimiento y prestar asesoramiento estratégico a las instituciones del Gobierno regional y también a los agentes sociales y económicos. El organismo acometió recientemente, en 2013, una reestructuración por la que se creó una división de investigación y evaluación, otra de información y datos y otra de prospectiva.

El sistema de evaluación en Bélgica

Valoración general

Bélgica es un Estado federal dividido en diferentes comunidades lingüísticas y culturales, una división que dota de una elevada complejidad a sus instituciones políticas, cuya organización y funcionamiento están condicionados por la necesidad de representar a esas diferentes comunidades. Esa representación se canaliza a través de unos partidos políticos que de ese modo penetran todas las instituciones de gobierno. La centralidad de los partidos políticos y la debilidad del Parlamento en su papel de órgano de control del Ejecutivo interfieren en el funcionamiento de

39. <http://www.iweeps.be/>

los mecanismos de rendición de cuentas y en la institucionalización de la evaluación. Este contexto es refractario a la introducción de factores técnicos que puedan contradecir posiciones de partido y poner en riesgo la estabilidad de la coalición en el gobierno. Por otra parte, el control de la iniciativa legislativa por parte del Gobierno minimiza la posibilidad de introducir disposiciones que favorezcan una evaluación independiente. En este entorno, el impulso de la evaluación parece depender directamente de la voluntad de los partidos en el gobierno. No obstante, su impulso se encuentra en gran parte ausente de los programas y de las campañas electorales.

Por lo tanto, la evaluación de políticas públicas continúa siendo algo raro, escasamente sistemático y disperso en el gobierno y en la Administración central belga, sobre todo debido a una toma de conciencia insuficiente sobre su utilidad. Sin embargo, en el plano sub estatal se han venido desarrollando dispositivos de evaluación desde mediados de 2000, con la región valona a la cabeza, a través del IWEPS (Instituto valón de la evaluación, la prospectiva y la estadística). No obstante, hay escasa información sobre el funcionamiento actual del sistema, tanto a nivel federal como en los estados que permita un análisis más detallado.

7. La evaluación en países de Europa meridional: Italia y España

Los países de Europa meridional han tratado de adoptar, sin conseguirlo del todo, tradiciones político-administrativas de raíz napoleónica (Ongaro, 2009; Pollitt i Buckaert, 2004). Tanto los elementos propios de esa tradición, como los factores contextuales, políticos y sociales, que han impedido su introducción plena en Europa meridional, han resultado en general poco propicios para el desarrollo de la evaluación.

Entre otros aspectos, en el modelo napoleónico de Administración Pública la concepción sobre las tareas a desarrollar por la Administración sigue siendo predominantemente legalista (versus “gerencialista” y discrecional); los valores que guían la provisión de servicios públicos priman la uniformidad (en lugar de estimular la diferencia y la experimentación), y los grupos sociales son percibidos como participantes no legítimos en el proceso de las políticas públicas. El Gobierno constituye de hecho la máxima autoridad en cuestiones de Administración y de control, que se ejerce a través de un sistema burocrático, regido por el Derecho Administrativo. El Parlamento, por su parte, pone en marcha comisiones parlamentarias para investigar determinados hechos, pero no hay una verdadera tradición parlamentaria de vigilancia (*oversight*) del ejecutivo. Los Tribunales de Cuentas son órganos judiciales, muy diferentes de los tribunales de cuentas o *Audit Offices* que se encuentran en países anglosajones o escandinavos.

Pero además, esos países comparten las rigideces típicas del modelo referido con las debilidades propias de una cultura social y política que no ha conseguido todavía erradicar la presencia e influencia de redes clientelares entre los partidos políticos, la administración y el sector económico y social (Ongaro, 2009; Kickert, 2011). De ello resulta a menudo una

excesiva penetración de la política en la Administración, que se identifica en la literatura como uno de los principales factores responsables del fracaso de los intentos de modernización de las administraciones en el Sur de Europa (Kickert, 2011).

Ello es así porque esa politización excesiva incide simultáneamente en tres dimensiones clave del funcionamiento de las administraciones y, por extensión, de las políticas que pretenden impulsar. En primer lugar, acentúa la orientación clientelar aludida, que favorece determinados intereses en detrimento del interés general. En segundo lugar, provoca el malbaratamiento de capital humano altamente cualificado y desincentiva la gestión y el desarrollo profesional de los recursos humanos. Finalmente, introduce una discontinuidad endémica en las políticas públicas, dado que cada gobierno tiende a imponer una nueva agenda, descartando de entrada o modificando substancialmente iniciativas que provengan de equipos anteriores.

Por otra parte, los sindicatos del sector público protegen la función pública y el empleo público en general tal como es, con sus defectos e ineficiencias, y se han opuesto explícitamente a la evaluación del rendimiento del personal al servicio de la Administración. Hay además, todavía, una elevada tolerancia ciudadana ante deficiencias o bajo rendimiento en la provisión de servicios públicos, aunque los recortes derivados de la actual crisis han actuado como revulsivo en algunos casos.

Todo ello ha constituido un contexto poco fértil para el desarrollo de procesos de reforma administrativa. Los progresos en esos ámbitos han tenido en todo caso un alcance limitado y fragmentado. Se trata de procesos discontinuos, localizados normalmente en organismos concretos, más que grandes reformas transversales.

La evaluación ha tendido a introducirse en ese entorno de forma exógena y reactiva, asociada a las obligaciones derivadas de la gestión de fondos estructurales europeos. Posteriormente han surgido procesos endógenos de institucionalización formal que no han llevado, al menos por el momento, al despliegue y funcionamiento de efectivo de auténticos sistemas de evaluación. Estos procesos conviven, sin embargo, como vimos en el caso de Bélgica, tanto en Italia como en España, con algunas iniciativas sub-estatales que, planteadas de forma más flexible, muestran una evolución interesante.

7.1. La evaluación en Italia

La evaluación se introdujo en Italia de la mano de factores externos y trató de implantarse, como decíamos, en un contexto político y administrativo extraño a esas prácticas.

Según Stame (2012), en ese país la penetración de los partidos políticos en todas las instituciones del Estado, incluida la Administración pública, y las consiguientes relaciones clientelares y uso instrumental de la administración al servicio de intereses partidistas, ha provocado una tradicional desconfianza y distanciamiento entre la sociedad italia-

na y el Estado. Ese distanciamiento ha impedido el desarrollo de una presión externa ciudadana y social suficiente sobre la Administración, a favor de una mayor efectividad, eficiencia y transparencia. Por el contrario, el Estado es percibido como un enemigo, a menudo con un excesivo celo controlador y represor, lo que agudiza la desconfianza y el alejamiento por parte de los ciudadanos.

En contraposición, existe una sociedad civil vital y organizada, que ha contribuido a la existencia de administraciones eficientes en el nivel local y comunitario, así como diversas ONG e iniciativas comunitarias que proveen servicios a colectivos necesitados. El papel de la sociedad civil organizada ha tomado más peso con la reciente crisis fiscal y la reducción de los presupuestos públicos.

En ese contexto, Ongaro y Valotti (2008, citados por Stame, 2012), caracterizan el sistema político-administrativo italiano como una mezcla discordante de instituciones centralizadas⁴⁰, tolerancia hacia la ineficiencia y valores democráticos. Esa falta de alineación provoca un debate político vibrante y es el contexto donde se han llevado a cabo diversas reformas de la Administración desde la década de los años 90.

Las prácticas

El ciclo de vida de las políticas públicas se contempla desde una perspectiva legalista y hay una aproximación eminentemente normativa a la implementación. Toda la atención se concentra en el proceso legislativo. Una vez la norma está aprobada, se sobreentiende que la implementación es algo prácticamente automático, sin consideraciones sobre diversas alternativas posibles, etc. En general, se da escasa importancia al proceso de implementación, y a la actuación de la Administración en general (en parte derivado del círculo vicioso al que aludíamos antes entre desconfianza, distanciamiento y desconocimiento), por lo que también son escasos los esfuerzos para procurar un mejor funcionamiento. Toda la atención, como decíamos, se concentra en el proceso legislativo y ello se traduce en intentar arreglar problemas de ineficiencia y mal funcionamiento a base de promulgar nuevas leyes. En palabras de Stame:

“...all the attention is turned on the legislative function, while the executive function (performed by the public administration) is considered less important. So, if it does not take place as expected, the tendency is not to address administrative problems but to promulgate a new law”. (Stame, op. cit., pos. 504)

Las reformas de la Administración Pública

Desde principios de los 90, se aprobaron una serie de leyes orientadas a la reforma de la administración pública, conocidas como “leyes Bassanini” por el Ministro que las impulsó, de gobiernos de centro-izquierda. Esas iniciativas incluyeron tanto reformas “tradicionales”

40. Sin embargo, aunque el sistema político es centralizado, pero ha dotado de un elevado nivel de autonomía a las regiones y a los municipios, y la reforma constitucional de 1999-2001 introdujo un sistema “cuasi federal” (el Estado, las regiones y los municipios se sitúan al mismo nivel como instituciones de la República).

del sector público (Wollmann, citado por Stame, 2012) –reforzar el liderazgo del Primer Ministro y de los alcaldes, descentralización de competencias a las regiones, transparencia, democratización, participación ciudadana...– y también otras inspiradas en el *New Public Management* –como la externalización en la provisión de servicios, la agencialización, la sujeción al derecho laboral de la contratación de directivos y gestores públicos...-.

Una reforma de especial importancia fue la orientada a delimitar más claramente los ámbitos de actuación política y administrativa o de gestión pública. Ello llevó a la contratación de directivos y gestores públicos provenientes del exterior de la Administración, a través de contratos sometidos al derecho privado, con duración limitada y que conllevaban la fijación de objetivos y el seguimiento y evaluación del grado de éxito alcanzado. Este tipo de vínculo comportó más autonomía en la gestión por parte de estos directivos y gestores, y también la creación de unidades de evaluación en todas las administraciones (orientadas a la evaluación del rendimiento o *Performance Measurement*).

Sin embargo, esas reformas han tenido un éxito limitado: los políticos han continuado influyendo de forma decisiva en el nombramiento de directivos y gestores públicos, que son removidos de sus cargos cuando los primeros cambian. Bassanini ofrece un balance elocuente y amargo:

“Accountability, merit, performance assessment are still the exception, not the rule; they have been overwhelmed by the spoils system, by the scattershot distribution of performance bonuses, by the explosion of the costs of politics, by the resistance of an entrenched bureaucratic culture, by the conservatism of the bodies responsible for auditing, by the inability of the political class to conceive of public policies in terms of strategies, objectives and quantifiable, measurable results”. (Bassanini, 2009, pág. 375)

En particular con relación a la evaluación del rendimiento, el mismo autor señala que las principales reformas, contenidas en el Decreto legislativo 80/1998, establecían que

“...for each administration the political authorities are responsible for defining public policies, translating them into strategic guidelines, setting precise, quantified goals, and introducing objective, reliable systems of evaluation; that administrators are completely autonomous and accountable for managing their units; that careers, promotions, removal from position, and even a part of pay must be related to results; and that public employees can be dismissed under the same rules that prevail in private law”.

En cuanto a sus resultados:

“...only a few administrations have applied them with conviction. The objectives, when they exist, are often generic and vague; performance is either not measured at all or is assessed in purely summary and discretionary fashion...” (Bassanini, 2009 págs. 379-381)

Respecto a las principales causas del fracaso, menciona la reticencia de la clase política italiana a fijar objetivos precisos de productividad y a desplegar mecanismos rigurosos y fiables para medir su consecución. Ello muestra la resistencia de la clase política a despojarse de amplios márgenes de discrecionalidad propios de una cultura clientelar. Por otra parte fijar objetivos relevantes y viables no es fácil; requiere tiempo, análisis y reflexión por parte de cargos políticos y de profesionales de la administración, tareas que están lejos de ser práctica habitual en ambos casos.

Contribuyó también la falta de continuidad de las políticas de reforma en los sucesivos cambios de gobierno, a pesar de tratarse de una reforma (la iniciada en los 90) pactada por el gobierno de centro-izquierda con la oposición de centro-derecha. También, la reforma fallida hacia un sistema federal, mencionada anteriormente⁴¹.

Un problema añadido fue la falta de recursos, ya desde el principio, ya que el estado de las finanzas públicas impidió realizar las inversiones necesarias para llevar adelante las reformas (por el contrario, un objetivo prioritario era la reducción del gasto público), incluyendo la informatización, el reconocimiento del mérito y de los buenos resultados, la mejora de la calidad de los servicios públicos, la formación del personal, la contratación de personal joven y/o experto en nuevas materias priorizadas en la reforma, etc. (Bassanini, op. cit.)

La reforma administrativa incluyó el establecimiento de unidades de evaluación en diferentes organismos públicos. Pero en numerosas ocasiones esas unidades fueron establecidas de forma un tanto precipitada, como si se tratase de una unidad administrativa más, replicando funciones de control de unidades anteriores en lugar de tomar en consideración cuáles deberían ser sus objetivos y cómo deberían cumplirlos (en particular, a través de qué métodos de trabajo, y qué perfiles profesionales serían necesarios para desarrollarlos e implementarlos) (Stame, 2012).

Sin una reflexión previa sobre estas cuestiones, las unidades de evaluación se constituyeron con personal generalista, proveniente de disciplinas diversas (juristas, contables, estadísticos...) y/o con especialistas en determinadas políticas sectoriales (salud, medio ambiente, etc.), sin que la experiencia en evaluación fuese un requisito. Como era de esperar, la forma de actuar de esas unidades no parece haber contribuido al avance de la evaluación.

Por otra parte, se ha hecho un gran esfuerzo académico (Borgonovi, 2005, citado por Stame, 2012) por establecer un sistema de medida de la productividad de las administraciones públicas, basado en baterías de indicadores de rendimiento (*Performance*

41. Fallida, según Bassanini (op.cit), por la ausencia de una cláusula de supremacía en la Constitución, a favor del gobierno central, que ha provocado duplicaciones en la legislación entre el Estado y las regiones, y ha impedido suprimir de la legislación estatal materias que deberían ser reguladas por la regiones, al no poder garantizar que esa regulación regional se haga de manera suficientemente homogénea como para salvaguardar el interés general de la nación y de los ciudadanos en todo el territorio del Estado.

Indicators), de inputs, outputs, outcomes... Pero esos indicadores se han concebido sin participación de los gestores y técnicos que deberán informarlos, lo que ha comportado limitaciones importantes en el diseño y utilidad real de los indicadores, así como en su apropiación –que ha devenido ritualización- por parte de la Administración:

“Unless the personnel are involved in their definition and in interpreting their results, indicators alone can favor goal displacement (typical of public sector ritualism); alternatively, if they measure a failure they may provoke disenchantment instead of a desire for understanding what specific conditions and human capacity may make a program succeed.” (Stame, op. cit., pos. 573)

Las funciones de evaluación han sido vistas, por tanto, como funciones de control, y los organismos públicos han desarrollado estrategias para sustraerse a su influencia, lo cual ha resultado contraproducente para el avance de la evaluación:

“...the implication was that since administrative personnel are familiar with control processes, they could better understand evaluation if it were presented as a further control; in reality, people are used not only to controls, but also to avoidance of them. This attitude has had negative consequences for evaluation.” (Stame, op. cit., pos. 555):

A finales de la década de 2000 se emprendió una nueva reforma, de la mano del Ministro Renato Brunetta, centrada de nuevo en aumentar la productividad de las administraciones públicas, a la vez que su eficiencia y transparencia. La reforma pivotó alrededor del concepto de mérito de los directivos, gestores, y del personal en general al servicio de las administraciones públicas, y su evaluación a través de un sistema de indicadores que hacían énfasis en la satisfacción de los usuarios de los servicios. Sin embargo, este nuevo intento de reforma fracasó, siendo uno de los motivos la dificultad para fijar objetivos:

“Setting challenging but credible objectives that can trigger significant improvements in the quality of services and efficient performance is a difficult task, requiring, upstream, an ability to elaborate and define the public policies for which the setting of objectives is an essential instrument, and a capacity for dialogue with the administrative units and their executives on the one hand and citizens/customers/users on the other. Without an adequate definition of the objectives, the whole reform mechanism will go nowhere...” (Bassanini, op. cit. pág. 389)

En el marco de la Reforma Brunetta se crea la *Commissione per la valutazione, la trasparenza e l'integrità delle amministrazioni pubbliche* (CIVIT), adscrita al Parlamento⁴². La CIVIT se creó por el Decreto Legislativo 150/2009, con la función, entre otras, de coordinar y supervisar el ejercicio independiente de las funciones de evaluación en

42. <http://bit.ly/1vTpcVv>

las administraciones, así como garantizar la transparencia de esas actividades. La CIVIT quedó absorbida en 2014 por la *Autorità Nazionale Anticorruzione*, sin haber desarrollado, por el momento, una actividad evaluadora relevante.

La evaluación de los fondos estructurales de la Unión Europea.

Las obligaciones de evaluación derivadas de la recepción de fondos estructurales de la UE provocaron la creación de unidades de evaluación (*Nuclei di Valutazioni*) en todas las regiones y ministerios con responsabilidad en la gestión de esos fondos, así como una red de evaluación (*Rete dei Nuclei*, creada por la ley 144/1999). Esas unidades de evaluación se han ubicado normalmente en departamentos de programación de gobiernos regionales y ministerios, y se han provisto de personal funcionario ya existente, o bien de personal experto contratado *ad hoc*, normalmente con perfiles de estadística o economía. Las evaluaciones normalmente se han encargado a expertos externos, mientras que esas unidades han gestionado los encargos y contratos correspondientes.

Dada la composición de esas unidades de evaluación, los estudios llevados a cabo no han sido propiamente evaluaciones, sino especialmente análisis económicos sobre el impacto de los fondos y de los programas ejecutados en la economía de la región y del país. Además, esas unidades son también responsables de proveer a la DG Regio de la Comisión Europea de un sistema de indicadores específico, la mayor parte de tipo financiero, sin lugar para el análisis de los resultados obtenidos y de cómo se ha llegado a ellos. Ese sistema de indicadores se implementa, de hecho, de forma centralizada en esas unidades, sin contacto con los beneficiarios ni con el personal a cargo de los programas. La propia DG Regio ha sido bastante crítica con las evaluaciones realizadas. Algunas regiones, como Campania, han sido capaces de desarrollar métodos más robustos y más relevantes a efectos de evaluación. Por su parte, UVAL (*Public Investment Evaluation Unit*, adscrita al Ministero dello Sviluppo Economico⁴³) desarrolla actividades de formación y de mejora de la capacidad de evaluación en el ámbito de los programas financiados con fondos estructurales.

Como conclusión, los rasgos característicos de la institucionalización de la evaluación en Italia pueden resumirse del modo siguiente (Stame, op. cit.):

- **Centralismo:**

"Evaluation units (institutions) are created from the above, they receive the templates for evaluation from the center. This influences the evaluation practice: rarely do the evaluation units work with their immediate evaluatee... This produces a resistance from the evaluatee in providing information and participating in the evaluation. Therefore, it is more difficult for the center to understand why something worked or did not..." (Stame, op. cit., pos. 817).

43. http://www.dps.mef.gov.it/uval_eng/presentation.asp

- **Legalismo:** lo que importa es que la evaluación esté regulada y que la ley se implemente. La utilidad de la evaluación para aprender y mejorar, y para rendir cuentas, queda en segundo plano.
- **Normativismo:** al equipararse con otra actividad administrativa más, se ha dado por descontado que sólo hay una forma correcta de realizar evaluaciones, sin poner en discusión métodos, etc. (aunque últimamente hay más debate al respecto, y se está dando entrada a métodos experimentales, quizás a causa de los cambios que ha habido al respecto en ciertos programas de la Comisión Europea). En su expresión probablemente más negativa, esto ha comportado la aceptación pasiva de sistemas de indicadores impuestos desde arriba, sin cuestionar su relevancia y su utilidad.

En palabras de Stame:

"All this has given rise to an evaluation culture that is characterized by lifeless institutions (evaluations units undistinguishable from existing administrative departments), values of uncertain status (transparency and independence declared but not really pursued) and malpractices of accomplishment (avoiding relevant choices, most notably in methodology)" (Stame, op. cit., pos. 809)

El sistema de evaluación en Italia

Valoración general

La evaluación empezó a introducirse en Italia en los años 90, a raíz del paquete de reforma administrativa conocido como "leyes Bassanini", que buscaban incrementar la eficiencia en la Administración a través, entre otras medidas, de la fijación de objetivos y la medición posterior del rendimiento. Aunque una parte importante de las medidas adoptadas tenía que ver con la gestión del personal al servicio de las administraciones, y por tanto con la valoración del rendimiento individual, en el marco de estas reformas se introdujeron también unidades de evaluación en diferentes organismos de la Administración. Estas unidades, sin embargo, no llegaron a cuajar. Una aproximación excesivamente legalista y "de arriba abajo", sin tener en cuenta las especificidades de la evaluación, sus exigencias ni tampoco el contexto en que se iba a introducir, abocó estas reformas al fracaso. Un fracaso al que contribuyó, además de los factores mencionados, la falta de continuidad del apoyo político a las reformas y de recursos suficientes. A finales de la década de 2000 vuelven a impulsarse medidas de reforma que incluyen la evaluación. Se crea la Commissione per la valutazione, la trasparenza e l'integrità delle amministrazioni pubbliche (CIVIT), adscrita al Parlamento. La CIVIT tiene entre otras funciones, la coordinación y supervisión del ejercicio independiente de las actividades de evaluación en las administraciones, así como garantizar la transparencia de esas actividades. Sin embargo, no parece haber desarrollado una actividad evaluadora relevante. En el nivel regional, las obligaciones de evaluación asociadas a la recepción de fondos estructurales han llevado a la creación de unidades de evaluación en diferentes regiones. No obstante, hasta el momento no parece haberse desarrollado tampoco a ese nivel un sistema en el que la evaluación esté plenamente asentada.

7.2. La evaluación de políticas públicas en España

A partir de la Constitución de 1978, España adoptó un modelo de Estado altamente descentralizado, que se ha ido implantando y profundizando en fases sucesivas y cuyo desarrollo ha copado gran parte de la agenda política, entrecruzándose con avance del Estado del Bienestar en ese mismo período (Gallego y Subirats, 2012). Los procesos de reforma de la Administración General del Estado acometidos en esos años han tenido en todo caso un alcance limitado, fragmentado, y centrado principalmente en aspectos de racionalización de las estructuras y de los procedimientos internos, así como en la profesionalización burocrática de la función pública. Un proceso caracterizado por cambios sucesivos, con efecto acumulativo y localizado normalmente en organismos concretos, más que grandes reformas transversales⁴⁴.

La práctica y el grado de institucionalización de la evaluación se encuentran entre los más bajos de la OCDE (Furubo y Sandahl, 2002, Varone et al. 2005, Viñas, 2009), registrándose algún progreso en la segunda mitad de la pasada década. En general, tal como afirman Feinstein y Zapico-Goñi (2010), más que un sistema de evaluación coherente, en España existe una constelación de organismos de diverso rango, que operan en distintos sectores de intervención pública, en diferentes niveles de gobierno, que tienen asignadas responsabilidades bien de evaluación, bien funciones relacionadas.

Los factores que explican esta situación son, entre otros (Viñas, 2009): el menor desarrollo del estado del bienestar y del volumen de inversión pública asociada; la débil tradición en investigación social y en su aplicación al estudio de las políticas públicas; el dominio de una cultura jurídico-administrativa entre las élites políticas y en el sector público, orientada fundamentalmente al control de la legalidad en detrimento de la efectividad; la falta de incentivos para impulsar la evaluación, tanto en el entorno político como en el administrativo; un sistema parlamentario estatal y autonómico sometido a la disciplina de partido, con limitadas posibilidades de control efectivo del ejecutivo.

Inducida fundamentalmente por las obligaciones asociadas a la recepción de fondos estructurales de la UE, la práctica de la evaluación se ha venido desarrollando a partir de los años 90, aunque de forma fragmentada y poco sistemática (Viñas, 2009). En algunos sectores como las políticas de salud⁴⁵ –en estrecho contacto con la cultura y la práctica científica– y las políticas de cooperación para el desarrollo⁴⁶ –expuestas a las

44. Como la reforma de la Agencia Española de Administración Tributaria, el Instituto Nacional de la Seguridad Social, el Servicio de Correos (Parrado, 2008) o, en Cataluña la reforma del Instituto Catalán de Salud (Gallego, 2013). Excepciones, aunque con resultados limitados, los constituyen, en el caso de la Administración General del Estado, la instauración del Presupuesto por Programas, en 1984, la aprobación de la Ley Orgánica de Funcionamiento de la Administración General del Estado (LOFAGE), en 1997, el Estatuto Básico del Empleado Público, en 2007 o la Ley de Agencias (2005).

45. Ver por ejemplo el Plan de Calidad del Sistema Nacional de Salud <http://bit.ly/1vrJSu5>.

46. Ver Argilés (2014) para una visión panorámica de la evolución de la evaluación de la política de cooperación internacional para el desarrollo en España. La evaluación en ese ámbito se impulsa desde mediados de los 90, y en particular desde 2005, con la creación de la Dirección General de Planificación y Evaluación de Políticas para el

prácticas generalizadas de evaluación de los donantes internacionales- la evaluación ha penetrado a un ritmo más rápido. En otros, como la educación primaria y secundaria, existen sistemas de monitoreo⁴⁷, así como en la educación superior sistemas de evaluación, certificación y acreditación de enseñanzas, profesorado e instituciones⁴⁸, aunque ninguno de ellos implica evaluación de las políticas públicas existentes en esos ámbitos. En muchos otros sectores, la práctica de la evaluación –especialmente si nos referimos a evaluaciones rigurosas de efectividad- es escasa o todavía inexistente.

La Agencia Española de Evaluación

A partir de 2005, diversos acontecimientos en el ámbito estatal y autonómico señalan la entrada de la evaluación en las agendas políticas y la voluntad de avanzar en su institucionalización.

Destaca entre ellos la creación de la Agencia Estatal de Evaluación de las Políticas Públicas y la Calidad de los Servicios, a través de la disposición adicional primera de la Ley 28/2006, de 18 de julio, de agencias estatales para la mejora de los servicios públicos. La Agencia, que iniciará actividades en 2007, se constituye a partir de la Dirección General de Inspección, Evaluación y Calidad de los Servicios del Ministerio de Administraciones públicas, integrando la mayor parte de su personal, sin experiencia en evaluación de políticas públicas, y bajo un régimen de función pública.

El proceso de creación de AEVAL incluyó una fase previa de propuestas, a partir de una comisión de expertos establecida a iniciativa del Ministerio de Administraciones Públicas y constituida por expertos en diferentes disciplinas relacionadas con la evaluación y en Administración Pública⁴⁹. La Comisión realizó un análisis exhaustivo de experiencias internacionales y presentó un documento de diagnóstico y recomendaciones. Finalmente, AEVAL se constituyó como una agencia estatal –adscrita inicialmente al Ministerio de Administraciones Públicas, y desde 2012, al Ministerio de Hacienda y Administraciones Públicas– que actúa en el marco de contratos de gestión de una vigencia de cuatro años, donde se definen objetivos y líneas de actuación para el período.

La misión de la Agencia es la promoción y realización de evaluaciones de las políticas y programas públicos cuya gestión corresponde a la Administración General del Estado, aunque se prevé también la posibilidad de evaluar políticas de ámbito autonómico a

Desarrollo (DGPOLDE). Sin embargo, ese impulso no ha llegado todavía a consolidarse en términos de recursos destinados a la evaluación, producción de evaluaciones de calidad y utilización de las mismas. A partir de 2008, el proceso de institucionalización experimenta un cierto estancamiento. Sin embargo, el énfasis en la evaluación del IV Plan Director de la Cooperación Española (2013-2016), así como la nueva política de evaluación que de él se deriva y que se aprobó a principios de 2013 parece abrir un nuevo ciclo favorable.

47. Nos referimos al Sistema Estatal de Indicadores de la Educación, que incluye indicadores de resultados educativos basados en pruebas estandarizadas de competencias básicas, competencias clave, abandono temprano, etc. Ese sistema es responsabilidad del Instituto Nacional de Evaluación Educativa, del Ministerio de Educación, Cultura y Deporte (<http://www.mecd.gob.es/inee/>)

48. Nos referimos al sistema desarrollado por la Agencia Nacional de Calidad y Acreditación (ANECA) (www.aneca.es).

49. El informe resultante de los trabajos de esa comisión está disponible en <http://bit.ly/1Ehr9uq>

través del establecimiento de convenios de colaboración con las comunidades autónomas. Por otra parte, es también misión de la Agencia promover la mejora de la calidad de los servicios públicos.

Como agencia ejecutiva adscrita al Gobierno, aunque con personalidad jurídica y presupuesto propio, AEVAL depende del Consejo de Ministros tanto en la designación de su presidente⁵⁰ como para la aprobación de los planes anuales de evaluaciones a realizar. Por otra parte, AEVAL se rige por un Consejo Rector que cuenta con representantes de los Ministerios de Hacienda y Administraciones Públicas (los Secretarios de Estado de Administraciones Públicas y de Presupuestos y Gastos), de la Presidencia del Gobierno (el Director General de Relaciones con las Cortes), del Ministerio de Asuntos Exteriores y Cooperación (el Subsecretario del Ministerio), y expertos independientes de reconocido prestigio. Asimismo, los representantes de las Comunidades Autónomas que celebren un convenio de colaboración con la Agencia podrán participar institucionalmente en su Consejo Rector.

El primer contrato de gestión suscrito entre AEVAL y el Gobierno, para el período 2008-2011, fijó como prioridad la evaluación del Plan Nacional de Reformas (PNR)⁵¹, aunque también podían incluirse en los planes anuales de evaluación otros programas a demanda de los diferentes ministerios. En ese marco, la Agencia ha llevado a cabo evaluaciones de políticas en sectores diversos, centradas sobre todo en el diseño y la implementación de las mismas y con enfoques metodológicos también distintos. Por otra parte, AEVAL publicó en 2010 una guía sobre fundamentos de evaluación de políticas públicas, como primer ítem de una colección pendiente de completar.

La actividad de evaluación de políticas públicas por parte de AEVAL se ha visto lastrada por diversos factores, ya desde su puesta en marcha. Entre los más importantes se encuentra seguramente la rigidez de la plantilla asignada a la Agencia, constituida por cerca de 60 funcionarios públicos de los cuales cerca de la mitad tienen encomendadas funciones de apoyo y gestión administrativa del propio organismo, y el resto se reparte entre los departamentos de evaluación y de calidad de los servicios. El origen de la Agencia como antigua Dirección General de Inspección, Evaluación y Calidad de los Servicios, junto al escaso margen de que ha disfrutado en la gestión de su plantilla, ha determinado hasta el momento una mayor especialización en el área de calidad de los servicios que en evaluación de políticas públicas.

Pero además, AEVAL ha debido desarrollar su actividad evaluadora con un apoyo más formal que real por parte de las instancias políticas. Si bien la creación de la Agencia fue en su día la expresión de una clara voluntad política de impulsar la evaluación, esa voluntad no ha tenido siempre continuidad en los sucesivos cambios ministeriales y

50. Desde su puesta en funcionamiento, hace ocho años, AEVAL ha tenido tres presidentes, dos de los funcionarios de carrera de perfil técnico, y uno de perfil político.

51. El PNR, adoptado en 2005, es un paquete de reformas alineadas con la Estrategia de Lisboa, que preveía la plena convergencia de indicadores de renta, empleo y ciencia con la media europea en 2010.

de gobierno. Tampoco se ha sabido generar a partir de ella una política específica de impulso de la evaluación en la Administración General del Estado que, más allá de crear instituciones formales a cargo de su ejecución, emprendiese también acciones de sensibilización, de generación de capacidad evaluadora en los diferentes ministerios y organismos públicos, de incentivación de la práctica evaluadora, de intercambio de experiencias y coordinación transversal de actividades de evaluación, etc.

Como balance, entre 2007 y 2013, la Agencia ha producido 36 informes de evaluación. Llama la atención –e ilustra bien la alusión anterior a la falta de continuidad en el apoyo político al impulso de la evaluación– la tendencia decreciente del número de informes publicados en el web de AEVAL, que ha pasado de 9 en 2007 a 3 en 2013, con una fuerte incidencia de las elecciones generales (en 2008, cuando disminuye un tercio el número de informes publicados, y en 2012, año en que no se publica ningún informe).

La evaluación en las Comunidades Autónomas

España es uno de los estados europeos con un nivel más elevado de descentralización territorial. Como vimos en apartados previos de este informe, los sistemas políticos altamente descentralizados ofrecen un contexto interesante para la evaluación, ya que configuran una especie de laboratorio natural donde las políticas se implementan en unos territorios y no en otros, o bien lo hacen a través de variantes, ofreciendo oportunidades para la construcción de escenarios contrafactuales que son más difíciles de generar en estados unitarios.

Sin embargo, esa oportunidad no se ha aprovechado suficientemente hasta la fecha. Como un intento en esa dirección cabe mencionar la constitución, en Noviembre de 2009 y en el marco de la Red Estatal Interadministrativa de Calidad de los Servicios Públicos impulsada por AEVAL, de un grupo de trabajo sobre evaluación de políticas públicas con participación de diversas comunidades autónomas, de la propia AEVAL y de la Federación Española de Municipios y Provincias (FEMP).

Ese grupo de trabajo llevará a cabo, a lo largo de 2010 y bajo la coordinación del Instituto Catalán de Evaluación de Políticas Públicas (Ivàlua), un intercambio intenso de información sobre diferentes aspectos de la situación y el impulso de la evaluación en las comunidades autónomas participantes: Castilla y León, Cataluña, Extremadura, Islas Baleares, Navarra, País Vasco, Valencia, además de AEVAL y la FEMP. Sin embargo, a partir de 2011 los recortes presupuestarios –que comportan la supresión de unidades de evaluación de reciente creación en algunas comunidades autónomas, la reducción de sus efectivos en otras, y la congelación de proyectos de implantación de nuevos dispositivos en otros casos– frenan las actividades del grupo.

No obstante, a pesar de la ralentización que ha supuesto la crisis económica en la institucionalización de la evaluación en las comunidades autónomas, es importante destacar algunas experiencias:

- A finales de 2005 se aprueba la Ley foral de evaluación de las políticas públicas y de la calidad de los servicios públicos de Navarra. Se trata de la primera ley que institucionaliza en España la evaluación de las políticas públicas como función transversal del gobierno y crea un órgano específico con esa finalidad: la Comisión Foral para la Evaluación de las Políticas y la Calidad de los Servicios. La Comisión es un órgano colegiado que tiene entre sus funciones el impulso de la cultura de la evaluación en la Comunidad Foral, la definición de criterios y metodologías comunes para su práctica, y la constitución de un registro de evaluaciones público, entre otras. Integra no solo diversos departamentos del Gobierno Foral, sino también las entidades locales y los agentes sociales, entre otros.

La gestión operativa del sistema se encarga al Instituto Navarro de Administración Pública, en el que se constituye un Servicio de Evaluación, con una dotación de dos personas, incluyendo la Jefe del Servicio. Ese servicio despliega una actividad intensa de sensibilización, de generación de recursos, de formación y de constitución de redes interadministrativas sobre evaluación en la Comunidad Foral. Sin embargo, con el advenimiento de la crisis económica el Servicio es suprimido.

- Un poco más tarde, en octubre de 2006, se constituye formalmente en Cataluña el Centro para la Investigación y Evaluación de Políticas Públicas –actualmente Instituto Catalán de Evaluación de Políticas Públicas (Ivàlua)-, bajo la forma de un consorcio público formado por las administraciones y las universidades catalanas. El Instituto inicia actividades en 2008 y, a pesar de un anclaje institucional algo débil, ha sabido afianzarse como referente de la evaluación en Cataluña. Lo ha hecho fundamentalmente a través de la realización de evaluaciones de impacto rigurosas –de políticas activas de empleo, sociales y, en menor medida, educativas– sometidas a la supervisión externa de expertos de reconocido prestigio. Por otra parte, Ivàlua ha desarrollado una estrategia de sensibilización y de capacitación de empleados públicos y también del tercer sector social que ha tenido una excelente acogida y que ha permitido conseguir avances significativos en la cultura de la evaluación en Cataluña. Sin embargo, la falta de una visión clara y compartida entre las élites políticas –más allá del nivel discursivo– sobre qué es la evaluación y cuáles son los componentes clave de rigor metodológico, de independencia y de transparencia que deben respetarse, pone en dificultades su avance.
- Posteriormente, en 2010 el Gobierno Vasco elabora el Plan de Innovación Pública 2011 – 2013, el cual incluye entre sus líneas estratégicas el fomento de la cultura de evaluación de políticas públicas. Ese Plan, y en particular la línea relativa a evaluación, es impulsado conjuntamente desde la Presidencia del Gobierno (*Lehendakaritza*) –en particular, desde la Dirección de Coordinación– y desde el Departamento de Justicia y Administración Pública, a través de la Dirección de Innovación y Administración Electrónica. Sin embargo, esa línea de trabajo parece diluirse con el cambio de gobierno, en 2012.
- Finalmente, en marzo de 2011 el Parlamento de las Islas Baleares aprobó la Ley 4/2011, de la buena administración y el buen gobierno. La norma crea una Oficina de Evaluación Pública, adscrita al Gobierno a través del Departamento competente

en calidad de los servicios. Confiere autonomía a los órganos del sector público balear para realizar las evaluaciones que estimen convenientes, aunque establece que el Consejo de Gobierno identificará anualmente una relación de intervenciones prioritarias, y la norma concede ese estatus de entrada a todas aquellas que superen un presupuesto de 5 millones de euros. La Oficina, y la política de evaluación en su conjunto contemplada en la norma, están, no obstante, pendientes de implantación.

Así, en España se han dado, desde mediados de la década de 2000, una serie de avances en la institucionalización de la evaluación, aunque sólo en algunos casos han dado lugar a una práctica evaluadora efectiva. La crisis económica y los recortes presupuestarios, sumados a la falta de una apuesta política clara por la evaluación, han supuesto una ralentización de ese proceso desde 2011.

El sistema de evaluación en España

Valoración general

La evaluación de políticas públicas en el seno de las administraciones se encuentra en una fase incipiente en España. Se introduce en un primer momento de forma exógena, vinculada a las evaluaciones preceptivas de fondos estructurales. Pero no será hasta mediados de la década de 2000 cuando aparezcan iniciativas de institucionalización endógena. Estas se materializan antes en la periferia del Estado (Navarra) que en la Administración central. En 2006 se crea la Agencia Española de Evaluación y Calidad, con la doble misión de evaluar las políticas públicas de la Administración General del Estado –y de aquellas comunidades autónomas que lo soliciten– y la calidad de los servicios públicos. La Agencia se crea a partir de una Dirección General de Inspección, Evaluación y Calidad de los Servicios del Ministerio de Administraciones públicas heredando su plantilla de inspectores y técnicos de gestión, y con escaso margen de maniobra para incorporar evaluadores expertos. Esta circunstancia, junto a la rotación en la Presidencia del organismo, y el tibio apoyo político de que ha disfrutado después de la primera etapa de puesta en marcha, han ido lastrando su desarrollo. Las actividades desarrolladas por AEVAL en el ámbito de la evaluación de políticas públicas se han centrado en la elaboración de informes, mientras que las acciones de sensibilización y de capacitación del personal de la Administración General del Estado han tenido menos peso y no parecen haber respondido a una estrategia clara. Por otra parte, el impulso dado a la evaluación por diversas comunidades autónomas desde 2005, entre las que destacan Navarra, Cataluña y el País Vasco, se frenó sustancialmente con el advenimiento de la crisis económica, paralizándose iniciativas y suprimiéndose dispositivos en algunos casos. Persiste en Cataluña el Instituto Catalán de Evaluación de Políticas Públicas (Ivàlua), que ha desplegado una estrategia de sensibilización, de capacitación y de evaluación rigurosa de impacto y se ha ido afianzando como referente de la evaluación en la comunidad autónoma y resto del Estado.

IV. La institucionalización de la evaluación en América Latina

De forma general, y a riesgo de simplificar en exceso, la trayectoria de la evaluación de políticas públicas en América Latina sigue una evolución más o menos paralela a la vista en Europa, en la cual se distinguen las siguientes fases (Neirotti, 2012):

- Una primera, que se alarga hasta finales de los 70, asociada a un desarrollo incipiente del Estado del Bienestar y caracterizada por las denominadas políticas desarrollistas, diseñadas bajo sistemas de planificación centralizada. Esta etapa se trunca en algunos países con el advenimiento de regímenes autoritarios (Chile y Uruguay, en 1973; Argentina, en 1976), aunque en otros es contemporánea a ellos (Brasil, Paraguay).

Aunque la evaluación de políticas públicas como tal todavía no aparece plenamente en escena, sí lo hace de forma incipiente bajo la forma de análisis *ex ante* de viabilidad económica, financiera y social de proyectos de inversión, con el apoyo de la Organización de Estados Americanos (OEA), el Banco Interamericano de Desarrollo (BID) y el Banco Mundial (BM) (Feinstein, 2012).

En esa etapa se generan nuevas fuentes de información necesarias para la gestión y el seguimiento de los planes de desarrollo (censos poblacionales, estadística económica, estadística educativa, etc.), que serán fundamentales a efectos de asegurar la evaluabilidad de las políticas.

- Una segunda fase, relacionada con la ola de reformas de la Nueva Gestión Pública, a partir de los 80, durante la cual se dan también los procesos de transición a la democracia en diversos países de la región (Chile en 1989, Argentina en 1983, Uruguay en 1984-89, Brasil en 1985-89, Paraguay, desde 1992...). La crisis de la deuda externa en la década de los 80 acentúa y amplía los esfuerzos realizados en la década precedente con relación a la evaluación *ex ante* y las unidades de seguimiento y evaluación se van propagando en la región.

En esta etapa se llevan a cabo grandes programas de ajuste estructural y sectorial. La evaluación como tal continúa relegada y el análisis *ex ante* de proyectos tiene también un peso importante. Así mismo, se ponen en marcha algunos dispositivos, como el Sistema Nacional de Evaluación de Resultados de la Gestión Pública (SINERGIA), en Colombia o, posteriormente, el Sistema Nacional de Evaluación (SINE) en

Costa Rica, ambos con apoyo del Banco Interamericano de Desarrollo (BID). También, dispositivos de evaluación asociados a los planes de actuación plurianuales del gobierno federal y de los municipios brasileños (Neirotti, 2012).

En Chile el Ministerio de Hacienda introduce a partir de 1993 indicadores de desempeño, y en 1996 la evaluación de programas del gobierno en base a revisiones del marco lógico. A finales de los 90, ese tipo de dispositivos se extienden a otros países de la región (Feinstein, 2012). En todos los casos se enfatiza el análisis de los resultados y de los impactos de las intervenciones públicas en el marco del proceso de modernización del Estado, aunque siguen una lógica no tanto evaluativa sino de *Performance measurement* y de valoración de la satisfacción de los usuarios de servicios públicos, propia de la Nueva Gestión Pública.

No obstante, durante esa etapa se da también una creciente profesionalización de la evaluación (Neirotti, 2012) y se estimulan acciones de “capacity building” tanto desde los diferentes gobiernos como por parte de organizaciones internacionales, aunque con más peso por parte de estas últimas y vinculadas a los programas de cooperación para el desarrollo en ámbitos estatales, regionales o locales, y en sectores como la educación, los sistemas de salud o los servicios sociales. Como vimos en el caso de diversos países europeos, también en estos años surgen en América Latina prácticas de comparación o *benchmarking* intermunicipal, regional o entre centros proveedores de servicios.

- Una tercera fase, que se inicia a finales del siglo pasado y que perdura en la actualidad, caracterizada por una “vuelta al Estado” –que retoma su rol de promoción económica y social aunque delegando algunas funciones en el sector privado-, acompañada de nuevas demandas de regeneración institucional y democrática por parte de una sociedad civil crecientemente movilizadora y organizada. En ese contexto, están surgiendo planteamientos de evaluación que enfatizan la participación de diferentes *stakeholders* e incorporar así diferentes perspectivas -no sólo referidas a diferentes niveles de gobierno implicados en las intervenciones evaluadas, sino también a actores sociales, grupos étnicos, etc.- y también con una finalidad de sensibilización y capacitación. Por otra parte, se van extendiendo los enfoques *mixed methods*, centrados en evaluar impactos o *outcomes*, pero capaces al mismo tiempo de entrar en la “caja negra” de la implementación para buscar en ella variables explicativas del grado de efectividad alcanzado.

En general, en los países latinoamericanos existe actualmente un amplio abanico de sistemas de evaluación, articulados a menudo alrededor de determinados sectores de intervención. Estos sistemas se han generado normalmente a partir de grandes programas cofinanciados por organizaciones en esos sectores, y con el tiempo han ido enraizando en sus respectivos entornos institucionales, a través, por ejemplo, de la creación de unidades de evaluación en organismos gubernamentales responsables de la gestión de esos programas. Esas estructuras han impulsado la demanda de evaluación y la profesionalización en este ámbito en los diferentes países, englobando normalmente diferentes disciplinas (sociología, antropología, economía, psicología,

salud, etc.). Ese proceso de profesionalización ha incluido la puesta en marcha de diversos programas de Master, así como la creación de sociedades de evaluación, también de dimensión transnacional, como la Red de Seguimiento, Evaluación y Sistematización de América Latina y el Caribe-ReLAC), la Red de Evaluadores de Naciones Unidas para Latinoamérica y el Caribe-EVALUN LAC), la Red Latinoamericana y del Caribe de Monitoreo y Evaluación-REDLACME) y la Plataforma Regional de Desarrollo de Capacidades en Evaluación y Sistematización de América Latina y el Caribe-PREVAL.

1. La evaluación en Chile

Chile cuenta con un sistema de monitoreo y evaluación de políticas públicas calificado a menudo de ejemplar (May et al., 2006). Ese sistema se ha desarrollado desde 1990⁵² en un contexto de estado unitario y un sistema político presidencialista, en el cual la iniciativa legislativa en materia de finanzas y presupuesto corresponde exclusivamente al Presidente. Ello significa que el Ejecutivo es responsable de la administración y gestión de las finanzas públicas, estando las competencias en la materia concentradas en el Ministerio de Hacienda.

En ese contexto, en Chile se ha desarrollado un sistema de monitoreo y evaluación complejo, vinculado a la gestión del presupuesto, impulsado “de arriba abajo” desde el Ministerio de Hacienda (MH) y más concretamente desde la División de Control de Gestión de la Dirección de Presupuestos (DIPRES). La Contraloría General de la República (equivalente a los tribunales de cuentas europeos) no ha jugado un rol equiparable al visto en algunos países europeos, ya que en Chile dicho órgano no tiene competencias en auditoría del desempeño.

Diversos factores han contribuido al diseño, puesta en funcionamiento y consolidación progresiva del sistema, en particular, la buena marcha de la economía chilena y de las finanzas públicas a principios de los 90, que estimuló el interés por lograr una aplicación más eficaz y transparente de los recursos públicos (Guzmán et al., 2014). En ese mismo sentido, durante esos años se llevan a cabo una serie de reformas en la gestión de las finanzas públicas que incluyeron la revisión del proceso de formulación y gestión del presupuesto. Y también otras iniciativas de modernización de la gestión pública, como la introducción de la planificación estratégica, el desarrollo de las TIC, instrumentos típicos del NPM como las cartas de servicios y, en general, la mejora de la atención a los usuarios. Así mismo, durante esos años se regula la Alta Dirección Pública, se potencia la formación de los funcionarios y se introducen sistemas de valoración del rendimiento vinculados a incentivos salariales.

52. Aunque los orígenes de la evaluación en Chile se remontan a los años 70, con la incorporación de evaluaciones *ex ante* a proyectos de inversión pública, en el marco del que hoy en día constituye el Sistema Nacional de Inversiones.

El sistema de monitoreo y evaluación cubre actualmente todos los organismos públicos del Gobierno y de la Administración central del Estado incluidos en la Ley de Presupuestos, y comprende diversos instrumentos de seguimiento, evaluación y control de gestión que han ido desplegándose y renovándose desde principios de los 90 hasta la actualidad a través de cuatro fases principales (Arenas y Berner, 2010; Guzmán, Irarrázabal y de los Ríos, 2014). Veamos a continuación esa evolución con algo más de detalle.

Primera fase: 1993 - 1999

En 1993, el Ministerio de Hacienda inició la implementación de mecanismos de seguimiento que permitieran generar información sobre el desempeño de programas y organizaciones públicas, tomando como referencia las iniciativas puestas en marcha en el Reino Unido bajo los gobiernos de Margaret Thatcher, vistas en apartados previos de este informe. La finalidad perseguida era introducir prácticas para mejorar la calidad del gasto y avanzar en materia de transparencia y rendición de cuentas (Arenas y Berner, 2010).

Para ello se implementó un proceso de planificación estratégica en cinco servicios públicos -de ámbito educativo, agrario, de previsión social, de tesorería y de hacienda-, a los que se asociaron sistemas de información para la gestión que permitieran la medición de indicadores de desempeño. En 1994 se extendió esta iniciativa a 26 instituciones del sector público chileno, poniendo en marcha un sistema de indicadores de desempeño y metas de gestión incorporados al proceso de formulación del proyecto de Ley de Presupuestos para 1995. El resultado de las metas de los indicadores era informado semestralmente por el Ejecutivo al Congreso Nacional. En 1997 la cobertura se incrementó a 67 servicios, representando cerca del 80% del total de instituciones susceptibles de integrarse en el sistema. El número de indicadores de desempeño prácticamente se triplicó llegando a 291 (Arenas y Berner, 2010).

Durante este período adquirió mayor relevancia el papel del Congreso Nacional en la discusión del Proyecto de Ley de Presupuestos, empezando a solicitar al Ejecutivo información acerca de los resultados de las instituciones públicas a través de la Comisión de Hacienda. En 1996 se firmó el primer Protocolo de Acuerdo entre el Congreso y el Ejecutivo con motivo de la tramitación del Proyecto de Ley de Presupuestos de 1997, que estableció una serie de compromisos tendentes a una mayor evaluación y transparencia de la gestión pública. Entre las iniciativas surgidas en este Protocolo destacan el compromiso de evaluar programas gubernamentales del área social, de fomento productivo y de desarrollo institucional y la obligación de que las instituciones públicas presenten anualmente balances financieros y de gestión.

Es así como en 1997 surgen dos nuevos instrumentos del sistema de monitoreo y evaluación:

- El Programa de Evaluación de Programas Gubernamentales (EPG), que en su fase piloto contempló la evaluación de 20 programas públicos.
- Los Balances de Gestión Integral, o BGI, que contenían las definiciones estratégicas de cada institución, resultados de algunos indicadores de desempeño y principalmente información de ejecución presupuestaria.

Es importante subrayar que en esta etapa el énfasis se pone en el diseño e implantación progresiva de los instrumentos, más que en la integración de la información resultante en la toma de decisiones. Ésta empezará a realizarse, también de forma gradual, sólo en la etapa posterior, a partir de 2000 (Guzmán et al., 2014).

Segunda fase: Implementación del Sistema de Evaluación y Control de Gestión (2000-2005)

En esta etapa se extiende la implementación de los instrumentos vistos en la etapa anterior en la Administración chilena, que se someten a algunos ajustes para su perfeccionamiento. En particular, se trata de mejorar las definiciones estratégicas contenidas en los BGI y los indicadores de desempeño como parte de la información que acompaña al envío del proyecto de Ley de Presupuestos al Congreso. Adicionalmente, se crearon nuevas líneas de evaluación, se incorporaron procedimientos para asegurar la difusión de los instrumentos y sus resultados, y su utilización en los procesos decisionales. Más concretamente:

- Las **definiciones estratégicas** contienen información sobre la misión, objetivos estratégicos, productos estratégicos (bienes y/o servicios) y clientes, usuarios o beneficiarios de los organismos públicos. Su incorporación en el proceso presupuestario tuvo por objeto incorporar información del por qué y cómo se iban a utilizar los recursos, contribuyendo a una discusión presupuestaria orientada a resultados. Desde 2000, las instituciones públicas deben informar anualmente sus Definiciones Estratégicas como parte del proceso de Formulación del Proyecto de Presupuesto de cada año.
- Las definiciones estratégicas son acompañadas por **indicadores de desempeño**, que pretenden medir el grado de progreso de cada institución hacia la consecución de las metas más relevantes que se han fijado. En concreto, los indicadores de desempeño permiten cuantificar y medir resultados en relación a los bienes y servicios que entrega o provee cada institución a sus beneficiarios, clientes o usuarios, monitorizando toda la “cadena de valor”: procesos, productos, resultados intermedios y resultados finales. Así mismo, deben capturar distintas dimensiones del desempeño: la eficiencia, la eficacia, la economía y la calidad. Por otra parte, los indicadores deben afectar a diferentes ámbitos del control, como son los procesos, los productos y los resultados, distinguiendo entre los resultados intermedios y los resultados finales. En esta etapa, sin embargo, la mayoría de indicadores estuvieron centrados en los niveles de producción de bienes y servicios, así como en aspectos de calidad, en

particular niveles de satisfacción de los usuarios (May et al., 2006). Los indicadores de impacto, en cambio, no estuvieron muy presentes en el sistema en esta fase, dada la mayor complejidad metodológica que requiere su definición y su medida, y también el hecho de que deben reservarse a programas que tengan una cierta madurez antes de empezar a medirlos.

Por otra parte, los indicadores de desempeño recogen información en el tiempo, por lo tanto las comparaciones se hacen principalmente respecto a desempeños pasados de la misma organización. La comparación de indicadores de desempeño entre programas públicos es, en esta fase, problemática porque todos ellos tienen diferentes objetivos, haciendo muy difícil encontrar indicadores homogéneos (May et al., 2006).

Las diferentes organizaciones públicas deben elaborar directamente los BGI –incluyendo las definiciones estratégicas y los indicadores de desempeño. La Dirección de Presupuestos del Ministerio de Hacienda tiene el rol de coordinar el proceso y de prestar asistencia técnica, así como de presentar posteriormente el sistema de indicadores al Congreso a efectos de rendición de cuenta, y utilizarla en los procesos de gestión y toma de decisiones sobre el presupuesto.

Sin embargo, durante esta etapa estos instrumentos se perciben más como instrumentos de control externo que como herramientas de gestión interna por las organizaciones públicas, por lo que no son plenamente “internalizados” por éstas. Así mismo, la disponibilidad y la calidad de la información necesaria para nutrir los diferentes elementos del sistema presenta márgenes de mejora no desdeñables (May et al., 2006).

- Por otra parte, en lo que se refiere a las **evaluaciones**, la línea de EPG, iniciada en 1997 y desplegada en esta etapa, se centra en análisis que parten del establecimiento de la matriz del marco lógico de la intervención para analizar después su justificación, su diseño, su organización y gestión, el grado de desempeño –definido en los mismos términos vistos para el sistema de indicadores–, y su sostenibilidad. Posteriormente, en 2001, se incorporan las evaluaciones de impacto y luego, en el año 2002, las evaluaciones institucionales, que evalúan un ministerio en su totalidad, incluido el conjunto de bienes y servicios que provee, y las evaluaciones comprensivas del gasto.

Las evaluaciones de impacto son en esta etapa de carácter retrospectivo y se realizan normalmente a través de métodos cuasi experimentales. Además, deben cumplir un conjunto de requisitos:

- En primer lugar, deben ser independientes, y para ello se establece que no las lleve a cabo la institución que tiene la responsabilidad de la ejecución del programa, sino paneles de expertos o empresas consultoras externas al sector público, contratadas y administradas por el Ministerio de Hacienda⁵³. Las instituciones participan

53. Las EPG se llevan a cabo a través de equipos de tres profesionales: un experto en evaluación, un experto en el sector de intervención del programa evaluado y un experto en gestión pública. Las evaluaciones de impacto se encargan a centros universitarios o empresas de consultoría. Finalmente, las EPN, que veremos en la etapa siguiente, se encargan normalmente a universidades, con el apoyo de un panel asesor de expertos internacionales que se crea *ad hoc* para cada evaluación.

durante todo el proceso de evaluación, pero la responsabilidad de la evaluación radica en el Ministerio de Hacienda como mandante de consultores y firmas consultoras externas.

- En segundo lugar, el programa de evaluación debe ser transparente, es decir que toda la información que proviene de las evaluaciones es de carácter público. Una vez finalizadas las evaluaciones, se envían al Congreso conjuntamente con una respuesta que efectúa cada una de las instituciones evaluadas y también con un resumen ejecutivo de ellas. Esta misma información es puesta en el sitio Web de la organización.
- Las evaluaciones deben cumplir también ciertos requisitos técnicos y para ello se desarrollaron un conjunto de documentos de apoyo, metodologías y procedimientos, incluyendo la participación de expertos como evaluadores y/o como asesores externos. Al respecto, el proceso de selección de evaluadores se lleva a cabo por concurso público y a través de la fijación de estándares elevados de calidad técnica.

Entre 1997 y 2004 se realizaron 158 EPG, que costaron en promedio el equivalente de US\$ 11.000 y tomaron entre cuatro y seis meses para completarse. Las evaluaciones de impacto más exhaustivas y rigurosas costaron en promedio US\$ 88.000 y se completaron en ocho meses; hasta 2004 se realizaron 14 de estas evaluaciones, es decir, alrededor de 4 por año. Análisis del sistema de evaluación desarrollado en esa etapa encontraron cierta disparidad en la calidad de ambos tipos de evaluación y apuntaron como posible causa los presupuestos estrictos y fechas de entrega ajustadas para estas evaluaciones. En ese sentido, algunos autores postulan la necesidad de revisar el modelo:

“¿Realizaría alguna vez el Ministerio de Hacienda de Chile una evaluación tan detallada y costosa como la de [el programa] Progresía [en México]? El peligro para Chile es que su énfasis en reducir al mínimo el costo de las evaluaciones quizá lo lleve a subinvertir en evaluaciones individuales, aun cuando los programas que se estén evaluando sean programas costosos de preponderancia política del Gobierno.” (May et al., 2006, pg. 22).

Como balance general de esta etapa (May et al. 2006):

- Se considera un rasgo muy positivo del sistema de monitoreo y evaluación su enfoque gradual, de modo que la instituciones puedan irse adaptando progresivamente a los nuevos instrumentos
- También se considera positiva la creación de áreas específicas de Control Administrativo dentro de los organismos públicos. Esas áreas son responsables de coordinar internamente el proceso de monitoreo evaluación, asegurar que la información necesaria esté disponible, etc. Sin embargo, se critica que los diferentes organismos no dispongan de un presupuesto específico para actuaciones de monitoreo y evaluación.

- Por otra parte, se señala como limitación la tentación de las instituciones de formular metas e indicadores muy básicos, que ya casi se han alcanzado, sobre todo si la evaluación del desempeño va vinculada a incentivos económicos para la institución y para los empleados. Se trata, por otra parte, de un problema común a sistemas de indicadores de desempeño (Van Thiel y Leeuw, 2002).
- Además, tanto las evaluaciones como el sistema de indicadores de desempeño se refieren a programas definidos en el ámbito de instituciones individuales, lo que no ofrece un buen encaje con una nueva generación de políticas sociales de naturaleza amplia y transversal, así como con intervenciones intersectoriales, donde la política implica diversas instituciones.
- Por otra parte, se señalan déficits en la apropiación del sistema y el aprendizaje institucional, propios de sistemas implantados “de arriba abajo”, además de una orientación excesiva al presupuesto y no tanto a resultados por parte del Ministerio de Hacienda. Sin embargo, existen mecanismos formales de *feed back* y de seguimiento de las recomendaciones contenidas en las evaluaciones, orientadas a facilitar el aprendizaje y su utilización en la mejora de los programas. Existen, de hecho, datos bastante precisos sobre el grado de utilización⁵⁴.
- Finalmente, se apunta también a una cobertura todavía baja en proporción al gasto público, así como a la falta de transparencia que comporta una difusión de la información en formatos aptos tan sólo para público experto, no adaptados al conjunto de la ciudadanía.

Tercera etapa: Consolidación y Fortalecimiento del Sistema de Evaluación y Control de Gestión (2006-2009)

En esa etapa se potenció el Sistema de Evaluación y Control de Gestión, incorporando medidas correctivas respecto a los déficits detectados en la etapa anterior, así como nuevos instrumentos (Arenas y Berner, 2010). En particular:

- Aumenta la cobertura de las evaluaciones, comprometiéndose la evaluación de 155 programas e instituciones entre 2006 y 2010, lo que involucra a uno de cada dos programas presupuestarios.
- Se crea de una nueva línea de evaluación en 2009, denominada Evaluación de Programas Nuevos (EPN). Esta línea es especialmente interesante, ya que se trata de evaluaciones prospectivas de impacto de los nuevos programas, basadas en métodos rigurosos de evaluación de impacto, preferentemente siguiendo diseños experimentales. El planteamiento prospectivo de la evaluación permite asegurar la disponibilidad de los datos necesarios para ese tipo de evaluaciones, imposible de obtener en muchos casos en evaluaciones retrospectivas.

54. Entre 2000 y el 2004, se determinó que el 25% de los programas necesitaban ajustes menores, por ejemplo con necesidades de cambios en la gestión interna, el monitoreo u otros. Por otro lado, el 10% de los programas evaluados fueron eliminados, mientras que un 21% de los programas sufrieron rediseños significativos de componentes y procesos de gestión interna. El resto comprendía modificaciones en el diseño del programa y en los procesos de gestión interna (39%) y la reubicación institucional del programa (5%). (Arenas y Berner, 2010).

- Por otra parte, se incluye la provisión de servicios de asistencia técnica a organismos públicos a partir de 2007, en relación con los diferentes instrumentos del sistema de monitoreo y evaluación.
- Se revisa el diseño y los sistemas de seguimiento de programas nuevos y de ampliaciones de programas presentados al Presupuesto desde 2008.

De forma complementaria:

- Se establece un Programa Marco de la Calidad incorporado a los Programas de Mejoramiento de la Gestión (PMG) en 2009, a través del cual las instituciones públicas deben certificar bajo norma ISO los procesos de provisión de bienes y servicios.
- Se crea, en 2008, el Fondo de Modernización de la Gestión Pública, con la finalidad de estimular la identificación y la solución de debilidades o problema de gestión de políticas, programas, instrumentos e instituciones públicas.

En general, según el análisis realizado por la Dirección de Presupuestos del Ministerio de Hacienda (Arenas y Berner, 2010), la aplicación del sistema en el período 2001-2010 muestra que ha aumentado la cobertura de los instrumentos y se ha elevado el nivel de exigencia hacia las instituciones respecto del cumplimiento de sus compromisos⁵⁵. En ese sentido, se estableció un sistema por el cual, una vez el presupuesto del ejercicio se ha aprobado, la DIPRES trabaja con cada organismo que ha sido objeto de evaluación, en la implantación de las recomendaciones de mejora. Estas se establecen en acuerdos formales o Compromisos Institucionales (Guzmán et al., 2014).

Cuarta etapa: 2010-2014

Desde 2010, el sistema de indicadores del desempeño se ha simplificado en cuanto al número de indicadores, y las evaluaciones han acentuado su orientación a medir el impacto, aligerando el análisis de la implementación. También se ha tratado de mejorar la utilización de la información producida por el sistema en la gestión del presupuesto. Además, han hecho su aparición nuevos actores: el Ministerio de Desarrollo

55. En 2000, 72 instituciones públicas comprometieron 275 indicadores, de los cuales sólo se evaluó el cumplimiento del 59%; en cambio, en la formulación del Presupuesto 2008, 142 instituciones comprometieron 1.443 indicadores, de los cuales fueron evaluados el 100%. También la calidad de los indicadores ha cambiado: de 1.684 indicadores comprometidos en 2004, 70% eran de producto o resultado; mientras que en el proyecto de Ley de Presupuestos 2010 un 20% de los indicadores miden resultados intermedios o finales y 71% resultados en el ámbito del producto. En lo que respecta a la evaluación, durante el período 1997-2010 se han evaluado 255 programas a través de la línea de EPG, 85 programas a través de evaluaciones de impacto y 44 instituciones a través de evaluaciones comprensivas del gasto. Además, se habían iniciado 9 evaluaciones de programas nuevos (EPN). En términos de recursos, entre 2000 y 2005 se había evaluado el equivalente a 12% del Presupuesto de ese último año. Considerando las evaluaciones comprometidas entre 2006 y 2010 el porcentaje equivale a 34% del Presupuesto de 2010. Finalmente, a través de los concursos del Fondo de Modernización de la Gestión se ha financiado la elaboración de 35 propuestas para solucionar una debilidad o problema de gestión de políticas, programas, instrumentos e instituciones públicas entre 2008 y 2009. Por otra parte, los recursos humanos destinados al impulso y coordinación del sistema de monitoreo y evaluación en el Ministerio de Hacienda experimentaron un incremento significativo a lo largo del período: en 2000, la División de Control de Gestión disponía de 11 profesionales, que aumentaron a 16 a finales de 2005 y a 27 en 2010.

Social (MDS), creado en 2011, y el Ministerio Secretaría General de la Presidencia (SEGPRES), con funciones de coordinación y seguimiento de las prioridades del Gobierno. Las funciones y recursos de la DIPRES no se han visto modificadas substancialmente, excepto en lo que respecta a las funciones de monitoreo y evaluación asumidas por el MDS. Éste ha asumido la realización de evaluaciones ex ante de programas sociales, así como su seguimiento. El MDS ha puesto en marcha un nuevo dispositivo: el Banco Integrado de Programas Sociales (BIPS)⁵⁶, que recoge información exhaustiva, accesible al público, sobre los programas sociales, incluyendo evaluaciones ex ante y retrospectivas.

Así pues, podemos concluir que Chile ha conseguido implantar en poco más de 20 años un sistema de monitoreo y evaluación con una amplia cobertura en la administración central del Estado. Ese sistema comprende, entre otros elementos, diversas líneas de evaluación de políticas públicas, incluyendo la evaluación rigurosa de impacto tanto retrospectiva como prospectiva.

Ello ha sucedido en el marco de una coyuntura favorable, de crecimiento sostenido de la economía y de estabilidad política y social. Y ha sido posible gracias a una combinación de voluntad política firme respecto al desarrollo de la evaluación -que ha tenido continuidad a través de los diferentes gobiernos- y a la existencia de capacidad técnica y recursos suficientes en el interior de la Administración, y en particular en el Ministerio de Hacienda. El sistema altamente centralizado ha contribuido, sin duda, en un principio, a facilitar la extensión del modelo, aunque esa aproximación centralizada se ha sabido combinar, especialmente a partir de mediados de la década de 2000, con incentivos y líneas de capacitación técnica que han facilitado su internacionalización.

El sistema de evaluación en Chile

Valoración general

Chile ha desplegado a lo largo de las últimas tres décadas un sistema de monitoreo y evaluación sofisticado y con una amplia cobertura. Los primeros pasos se sitúan a principios de los 90 y siguen el modelo de las reformas administrativas del Gobierno británico en los 80, con énfasis en el Performance measurement y la conexión con el presupuesto. Se trata de un proceso fuertemente centralizado en el Ministerio de Finanzas y, dentro de éste, en la Dirección General de Presupuestos (DIPRES) que sigue una lógica normativa y “de arriba abajo”, especialmente en las primeras fases. El Congreso aprovecha casi desde el principio y de forma creciente las oportunidades que el nuevo sistema ofrece para mejorar sus funciones de vigilancia y control del Ejecutivo. En 1997 se añade al sistema el Programa de Evaluaciones Gubernamentales (PEG), con un enfoque de evaluaciones “de escritorio” que siguen el modelo de análisis del marco lógico. Más adelante, en 2001, se incorpora un componente de

56. <http://www.programassociales.cl/>

evaluaciones de impacto retrospectivas, que se contratan a expertos externos y que irán ampliando su cobertura a lo largo de la década. En 2009 se añaden las Evaluaciones de Programas Nuevos, que son evaluaciones de impacto que se diseñan de forma prospectiva, desde el inicio de los programas, aplicando en lo posible métodos experimentales. El elevado grado de centralización del sistema y la aproximación “de arriba abajo” provocan al principio problemas de falta de apropiación y déficits de calidad en la información aportada por parte de los organismos públicos. La creación de unidades específicas responsables del sistema en cada organismo público, junto con actividades de asistencia técnica y formación introducidas a partir de la segunda mitad de la década de 2000, favorecen una mayor internalización. Por otra parte, el sistema de indicadores se ha ido perfeccionando, reduciendo el número y aumentando la orientación a medir resultados e impactos. Los enfoques y métodos de evaluación se han ido también refinando, priorizando la evaluación rigurosa de impacto. También ha mejorado la utilización de la información producida en la mejora de los programas, a través de compromisos formales y de un seguimiento sistemático. Al final del período se han incorporado nuevos actores: el Ministerio de Desarrollo Social, en lo que respecta a la evaluación de programas sociales, y el Ministerio de la Secretaría General de la Presidencia, con funciones de coordinación y seguimiento de las prioridades del Gobierno.

2. La evaluación en México

La práctica de la evaluación de políticas públicas empieza a sistematizarse en México a finales de los años noventa, coincidiendo con el inicio de un proceso de cambio político que permitió al país realizar avances substanciales en materia de calidad democrática, así como llevar a cabo diversos procesos de reforma administrativa.

México es una república federal de corte presidencialista, compuesta de 31 estados más el Distrito Federal. El poder ejecutivo es ejercido por el Presidente de la Nación, que es elegido mediante sufragio universal directo cada seis años sin posibilidad de reelección. El poder legislativo reside en el Congreso de la Unión, un órgano bicameral compuesto por la Cámara de Diputados y el Senado. Los senadores son elegidos cada seis años y los diputados cada tres, en ambos casos sin posibilidad de reelección inmediata. La organización política de los estados reproduce el mismo esquema, con elección directa de gobernadores cada seis años y elecciones legislativas cada tres. A pesar de la estructura federal, el proceso de toma de decisiones está altamente centralizado, tanto en la Federación como respecto a los Estados y dentro de éstos. Los liderazgos personales son muy importantes y la movilidad de cargos, bastante fluida, lleva a discontinuidades en la política institucional.

En las elecciones legislativas de 1997, por primera vez en la historia del México moderno el Partido Revolucionario Institucional (PRI) no obtuvo mayoría. Posteriormente, en

las elecciones presidenciales de 2000, el cambio político se consolidó con la elección de Vicente Fox, candidato del Partido de Acción Nacional. El PAN continuaría en el poder tras las elecciones presidenciales de 2006, en las que resultó elegido Presidente Felipe Calderón.

En los años anteriores a ese período de transformación política e institucional, la información disponible sobre los programas sociales tenía que ver sobre todo con la ejecución de sus presupuestos. No existía presión sobre el Ejecutivo por parte de otros poderes del Estado, ni tampoco exigencia ciudadana por conocer los resultados de las intervenciones públicas. Entre mediados de los años 70 y principios de los 90, hubo diversos intentos de crear esquemas de evaluación de los programas federales, generalmente asociados a la Secretaría de Programación y Presupuesto, que no progresaron por falta de consenso y sobre todo por la referida ausencia de una presión política y ciudadana suficiente.

Primera etapa: hacia un Sistema de monitoreo y evaluación basado en resultados (1999–2006)

A mediados de los años 90 se promovió una estructura de presupuesto por programas que sigue vigente en la actualidad y que supuso, entre otros efectos, una mayor implicación de la Cámara de Diputados en el seguimiento y control de las actuaciones del Gobierno Federal. Por otra parte, ya en la segunda mitad de los años 90, surgieron evaluaciones aisladas de programas sociales (los programas de Abasto Social de Leche, de Abasto Rural, de Capacitación laboral) que generaron información parcial sobre la situación de algún elemento de dichos programas.

Pero el hito que revolucionó la práctica evaluadora del país fue el conjunto de evaluaciones cuantitativas y cualitativas que se realizaron para el programa de transferencias condicionadas *Programa de Educación, Salud y Alimentación* (PROGRESA). En 1997, el gobierno Mexicano comienza una prueba piloto del programa, muy novedoso en aquellos años: se pagará a las madres una cantidad de dinero todos los meses a condición de que sus niños menores están al corriente de vacunas y cuidados médicos preventivos y si los niños mayores van al colegio. Debido a la falta de recursos y a complicaciones logísticas, la política no podía llegar a todos los pueblos del país a la vez. El conjunto de pueblos donde primero se implantó la política fue escogido al azar, lo cual permitió que la evaluación siguiese un diseño experimental. El nivel de rigor de la evaluación, su independencia y la valiosa información que aportó sobre el programa, la mantuvieron a salvo de manipulaciones políticas y PROGRESA sobrevivió al cambio sexenal ocurrido en 2000. El ejemplo de PROGRESA –que posteriormente se llamó OPORTUNIDADES y actualmente PROSPERA– promovió la expansión de evaluaciones de políticas públicas tanto en México como en buena parte del resto del continente americano.

Aunque quizás el cambio institucional más importante para avanzar hacia una cultura de evaluación en México fue la Ley aprobada en 1999 por el Congreso de la Unión, según la cual el gobierno mexicano debía presentar evaluaciones externas anuales de todos sus programas federales sujetos a reglas de operación, es decir programas de subvenciones. Este cambio repentino – pasando de prácticamente no tener evaluaciones a tener evaluaciones en cientos de programas anualmente– si bien implicó en la práctica un proceso lento de aprendizaje e implantación progresiva, contribuyó a familiarizar con cierta rapidez a numerosos funcionarios de la Federación con la evaluación. Así, en 2001 casi 100 programas sujetos a reglas de operación estuvieron obligados a tener una evaluación externa, realizada por instituciones académicas, de investigación, expertas en distintos temas, y estas evaluaciones fueron presentadas anualmente al Congreso para alimentar decisiones presupuestarias.

Destaca también en ese período de finales de los 90, y gracias también al impulso de la mayoría opositora en el Congreso, la reforma del marco institucional de la función de fiscalización superior del Estado. Esa reforma incluyó la modificación de diversos artículos de la Constitución en 1999 y la promulgación de la Ley de Fiscalización Superior de la Federación en Diciembre de 2000. En ese nuevo marco jurídico se creó la Auditoría Superior de la Federación (ASF), adscrita al Legislativo y en sustitución de la Contaduría Mayor de Hacienda. La ASF incluye entre sus funciones las llamadas Auditorías de desempeño, que se orientan a evaluar el grado de cumplimiento de metas y objetivos de los programas gubernamentales, siguiendo criterios de eficacia, eficiencia y economía, así como su impacto, –social y económico– y beneficios para la ciudadanía⁵⁷.

Casi al mismo tiempo, en 2001, un órgano de fiscalización interno del Ejecutivo, la Secretaría de la Contraloría y Desarrollo Administrativo (SECODAM) cambió de nombre y de atribuciones y pasó a llamarse Secretaría de la Función Pública (SFP). La SFP continuó con sus labores fiscalizadoras y de control respecto a las dependencias e instituciones federales⁵⁸, pero se concentró también en el impulso de diversas líneas de reforma administrativa.

57. En el web de la ASF (www.asf.gob.mx) se publican únicamente los llamados “Informes de cuenta pública” y otros informes relacionados con actividades de fiscalización, correspondientes a cada ejercicio. Sin embargo, la información e informes consultados al respecto en el web de la ASF, de carácter sintético, no ha permitido comprobar el alcance real de esas actuaciones.

58. La SFP cuenta con diferentes unidades con funciones de evaluación, auditoría y control gubernamental :

- Una Unidad de Evaluación de la Gestión y el Desempeño Gubernamental, que tiene entre sus funciones (art. 24): (a) Establecer, organizar y coordinar, “el sistema de control y evaluación gubernamental”. (b) Realizar las acciones que en el ámbito de competencia de la SFP “se requieran para el funcionamiento del sistema de evaluación del desempeño previsto en la Ley Federal de Presupuesto y Responsabilidad Hacendaria, a fin de conocer los resultados de la aplicación de los recursos públicos federales, el impacto social de los programas y proyectos realizados con recursos públicos e identificar la eficiencia, economía, eficacia y calidad en las dependencias, las entidades y la Procuraduría”. (c) Proponer “normas de carácter general, lineamientos y políticas que rijan el funcionamiento del sistema de control y evaluación gubernamental”. (d) Validar los indicadores de gestión que en términos de la Ley Federal de Presupuesto y Responsabilidad Hacendaria son la base para evaluar el grado de cumplimiento de los objetivos y metas de los programas y proyectos de las dependencias, las entidades y la Procuraduría, y el desempeño de éstas en la ejecución de los mismos. (e) Establecer criterios de evaluación, así como metodologías o modelos en esa materia que permitan la operación del sistema de control y evaluación gubernamental. (f) Administrar la información que en materia de evaluación se genere; realizar el análisis de la misma y formular recomendaciones.

En particular, en los ámbitos de transparencia, mejora regulatoria y simplificación administrativa, gobierno electrónico y profesionalización de la Función Pública (especialmente a través del establecimiento, en 2003, del Servicio Profesional de Carrera).

Por otra parte, también en 2001, en el ámbito de las políticas sociales, la Secretaría de Desarrollo Social (SEDESOL) estableció un área especializada compuesta por un equipo de profesionales con amplias capacidades técnicas en el área de evaluación, lo que otorgó un carácter formal y permanente a la función evaluadora en ese ámbito de política pública.

En 2004 se da un paso más a través de la aprobación de la Ley General de Desarrollo Social (LGDS), con un voto prácticamente unánime de todos los partidos políticos. La Ley busca institucionalizar diversas prácticas en las políticas sociales, entre ellas la de evaluación. En Agosto de 2005, la SEDESOL publica en el Diario Oficial de la Federación el Decreto para la creación del Consejo Nacional de Evaluación de la Política de Desarrollo Social (CONEVAL) como un organismo público descentralizado de la Administración Federal con autonomía técnica y de gestión. Asimismo, dicho pronunciamiento establece que sus objetivos son:

«Normar y coordinar la evaluación de la Política Nacional de Desarrollo Social y las políticas, programas y acciones que ejecuten las dependencias públicas, sin perjuicio de las atribuciones que en materia de control y evaluación tienen las secretarías de Hacienda y Crédito Público y de la Función Pública, y Establecer los lineamientos y criterios para la definición, identificación y medición de la pobreza, garantizando la transparencia, objetividad y rigor técnico en dicha actividad.» (Art. 3)

La Ley crea al CONEVAL como una institución con autonomía técnica y de gestión que tiene como objeto evaluar la política de desarrollo social, sus programas, fondos y acciones, así como medir la pobreza a nivel nacional, estatal y municipal. La institución tiene la atribución de evaluar –y coordinar la evaluación– de programas y políticas en el ámbito social en todas las secretarías del gobierno federal.

De acuerdo con la Ley General de Desarrollo Social (2004), el Consejo está presidido por el titular de la SEDESOL, seis investigadores académicos –elegidos a través de un proceso que incluye concurso público y el voto de los 32 gobiernos de las Entidades Federativas, la Cámara de Diputados, la Cámara de Senadores, representantes de los Municipios y el Gobierno Federal. También son miembros del Consejo un representante de la SHCP y dos integrantes del Órgano de Vigilancia. El peso que tienen en el gobierno

-
- Una Unidad de Control de la Gestión Pública, que tiene, entre otras atribuciones (art. 25) “establecer, organizar y coordinar, con la participación de la Unidad de Evaluación de la Gestión y el Desempeño Gubernamental, el sistema de control y evaluación gubernamental.”
 - Una Unidad de Auditoría Gubernamental, a cargo de un programa anual de auditorías y visitas de inspección a las entidades de la Federación.

de la institución los seis investigadores académicos, le confieren independencia técnica y refuerzan la credibilidad de sus actuaciones.

El Consejo consigue también un nivel razonable de independencia económica, a través de una dotación presupuestaria que es considerada suficiente para impulsar las evaluaciones de programas sociales que tiene encomendadas, al menos en esos primeros años (González de Licona, 2010).

Como balance general de la etapa, entre 2000 y 2006 se llevaron a cabo cerca de 500 evaluaciones externas, que fueron dirigidas al Congreso, a la SHCP, a la SFP y los propios gestores de los programas. Estas evaluaciones, en general, también se hicieron públicas. Sin embargo su utilización en la mejora de la política pública fue, según algunos protagonistas del proceso, prácticamente nula (González de Licona, 2010).

Segunda etapa: consolidación y expansión del sistema de monitoreo y evaluación (2007-actualidad)

La Ley de Presupuesto y Responsabilidad Hacendaria (LPRH) promulgada en 2006 y reformada en 2007, crea el Sistema de Evaluación para el Desempeño (SED), que tiene como objetivo realizar una valoración del desempeño de los programas y proyectos públicos para contribuir a un Presupuesto basado en resultados (PBR). En la práctica este sistema, desarrollado por la Secretaría de Hacienda y Crédito Público (SHCP), se constituye como el conjunto de los indicadores, evaluaciones y metodologías para tener información objetiva que ayude a tomar decisiones presupuestarias. El SED es de aplicación Federal, afectando a todos los programas públicos financiados a ese nivel, incluyendo los sociales. Pero la Ley también prevé que los estados y municipios avancen en materia de evaluación y de monitoreo.

En el Decreto de Presupuesto de Egresos de la Federación que el Congreso publicó en enero de 2007, se establecía que la coordinación de la evaluación de programas federales estaría a cargo del CONEVAL, la SHCP y la SFP. Para ello, decía el Decreto, se emitirían Lineamientos Generales para la Evaluación de Programas Federales. Esos Lineamientos se publican en marzo de 2007, con el objetivo explícito de regular el sistema de Monitoreo y Evaluación de los programas federales, a efectos de sistematizar el tipo de actividades a llevar a cabo y contribuir a la consolidación del SED. Así pues, esas reformas se orientaron a avanzar hacia el establecimiento de un vínculo entre la planificación, el presupuesto y la ejecución de los programas federales haciendo uso de herramientas de monitoreo y evaluación.

Por otra parte, el Plan Nacional de Desarrollo (PND) 2007-2012, que la Ley de Planeación pide al inicio de cada sexenio y que coordina SHCP y la Presidencia, incluyó dos elementos novedosos:

- Metas claras, cuantificables y, en muchas de ellas, indicadores de resultados (estratégicos) que reflejaran mejoras en el bienestar de la población.
- Se buscó la máxima correspondencia entre el PND y los diversos programas sectoriales, para los que también se definieron metas cuantificables e indicadores de resultados estratégicos. Ese mismo sistema de objetivos e indicadores estratégicos se estableció también para cada Secretaría o dependencia de la Federación. Esa fijación de metas e indicadores se llevó a cabo a través de una estrecha colaboración entre las Secretarías y la Presidencia de la República.

El PND 2013-2018, por su parte, mantendrá y reforzará esos elementos de monitoreo y evaluación, así como la coordinación e integración entre diferentes niveles de planeación (CONEVAL, 2014).

En cuanto a los elementos concretos del sistema de monitoreo y evaluación que se estableció a partir de 2007, a través de los Lineamientos Generales para la Evaluación de Programas Federales se impulsaron los siguientes:

- En primer lugar, para mejorar la lógica interna de los programas y que fuera más claro cuál es el fin último de los mismos, así como para definir indicadores de resultados, se pidió que todos los programas federales tuvieran una Matriz de Marco Lógico (MML), que en este contexto se llamó Matriz de Indicadores. Esa matriz servía también para explicitar la conexión entre los diferentes programas y los objetivos nacionales expresados en el PND.
- Por otra parte, para lograr una evaluación más coordinada y más homogénea de los programas y políticas del gobierno federal, los Lineamientos plantearon un Programa Anual de Evaluación (PAE), que incluyera las evaluaciones más importantes a realizar cada año.
- Además, en vez de tener sólo un tipo de evaluación, como el que se pedía desde 2000, los Lineamientos establecen diferentes instrumentos con el objetivo de hacer más flexible el esquema global de evaluación. En particular:
 - Evaluación de Consistencia y Resultados: analiza sistemáticamente el diseño y el desempeño global de los programas, sobre la base de la matriz de indicadores.
 - Evaluación de indicadores: analiza mediante trabajo de campo la pertinencia y alcance de los indicadores de un programa.
 - Evaluación de Procesos (o de implementación), para analizar la eficacia y efectividad de los procesos operativos para la mejora de la gestión.
 - Evaluación de Impacto, para medir el grado de efectividad alcanzado y atribuible al programa, utilizando grupos de control para separar la contribución de otros agentes en la consecución de los resultados finales.
 - Evaluación Estratégica, centrada en un programa o conjunto de programas entorno a una estrategia, política o institución.
 - Evaluación Específica, que son llevadas a cabo por las propias dependencias o instituciones, de forma voluntaria, para analizar algún aspecto preciso de sus programas o de su gestión.

En la etapa previa, las evaluaciones que pedía el Congreso no eran homogéneas y por tanto la información no era útil para analizar diversos elementos comparativos entre programas. Los Lineamientos buscan que en las evaluaciones existan elementos homogéneos para mejorar el control de calidad de las evaluaciones y para poder tener elementos de comparabilidad entre programas.

Según el Informe de Evaluación de la Política de Desarrollo Social 2014 (CONEVAL, 2014), entre 2007 y Junio de 2014 se llevaron a cabo 1.504 evaluaciones de programas, acciones y políticas de desarrollo social, así como informes periódicos de indicadores trimestrales cuyos resultados son publicados regularmente en el web de CONEVAL.

En esta etapa se introducen también mecanismos para potenciar el uso de la información generada a través del sistema de monitoreo y evaluación en la mejora de los programas y de las políticas evaluadas. Así, CONEVAL, la SHCP y la SFP emitieron en 2008 los Lineamientos para la Mejora Continua del Desempeño público. Este instrumento busca que a partir de los hallazgos encontrados por las evaluaciones, así como por otros instrumentos generados por la SFP y la SHCP, se elabore un plan de acción específico de mejora del desempeño. El instrumento no sólo contiene compromisos a cumplir por los gestores directos del programa, sino que también involucra a la institución en el que éstos operan y a otros actores externos del ejecutivo –la propia SHCP, otras dependencias-, que deben actuar conjuntamente para mejorar el desempeño de la política pública en la que se inserta el programa.

A partir de ese momento, CONEVAL realizó el seguimiento de las modificaciones en los programas a partir de las evaluaciones. Según datos del Consejo (CONEVAL), 2014, desde 2010 se han hecho modificaciones importantes en los programas sociales. Los datos más actuales muestran que en 2013-2014 se reorientaron sustancialmente 47 programas (el 50%), mientras que 39 corrigieron actividades o aspectos de su funcionamiento (el 41%). Entre 2010 y 2014, sólo un programa fue suspendido como consecuencia del resultado de una evaluación.

El cambio en la cultura de la evaluación también parece haberse ido extendiendo a las entidades federativas. Si bien ha sido históricamente más lento y desigual entre entidades, CONEVAL cuenta con un sistema de análisis del funcionamiento de los sistemas de evaluación y monitoreo en los estados, que manifiestan un sustancial avance en los últimos cuatro años (CONEVAL, 2013).

Principales retos del proceso de institucionalización de la evaluación

Se identifican como principales retos los siguientes:

- Las debilidades persistentes en la planeación, especialmente en el ámbito de las políticas sociales, a pesar de las mejoras registradas:

“...los objetivos de la política de desarrollo social se encuentran definidos de forma fragmentaria y dispersa; las capacidades básicas a las que un individuo debería tener acceso no se encuentran apropiadamente definidas en la legislación vigente, ya sea en la Constitución o en las leyes secundarias.” (CONEVAL, 2014, pg. 101)

- Las carencias todavía existentes -a pesar de los avances registrados- en la integración y coordinación entre los diferentes actores que forman parte del sistema de planeación, monitoreo y evaluación, y en particular de la Presidencia, la SHCP, la SFP y CONEVAL, todas ellas con competencias de impulso, supervisión y prescripción en materia de monitoreo y evaluación del desempeño.
- Asimismo, es importante cuidar la generación de información para asegurar óptimas condiciones de evaluabilidad. Al respecto, es importante fortalecer la capacidad técnica de los profesionales a cargo de los programas y de las áreas de evaluación y planeación en las dependencias federales; mejorar las capacidades internas para analizar la información básica del programa, no solo la presupuestal; continuar con las líneas de actuación ya existentes para la mejora de los registros administrativos para que apoyen el monitoreo y evaluación externa, y para la homogeneización de dichos registros administrativos, en estrecha colaboración con el Instituto Nacional de Estadística y Geografía, aprovechando para ello los comités de estadística y generación de información que coordina el INEGI en las diferentes dependencias y entidades públicas.
- Finalmente, hay que seguir impulsando el uso de las evaluaciones no sólo en la mejora de los diferentes programas y políticas, sino también para la mejora de la planeación general (en cuanto a la revisión de objetivos e indicadores de seguimiento), así como en la definición de los programas presupuestarios y en la gestión del presupuesto. mismas (para planear, diseñar, presupuestar, entre otros usos).

El sistema de evaluación en México

Valoración general

El desarrollo de un sistema de monitoreo de políticas públicas en México arranca a mediados de los años 90 a partir de la introducción del presupuesto por programas por parte de la Secretaría de Hacienda y Crédito Público (SHCP). En la segunda mitad de los años 90, el éxito de la evaluación del programa PROGRESA supone un punto de inflexión. En 1999 el Congreso aprueba una ley que obliga al Gobierno federal a presentar evaluaciones externas anuales de todos sus programas de subvenciones. Entre 2000 y 2006 se llevaron a cabo cerca de 500 evaluaciones externas. Aunque su utilización en la mejora de los programas fue marginal, esa etapa contribuyó extender la cultura de la evaluación en el sector público. En paralelo, en 2001 hace su aparición en el sistema la Secretaría de la Función Pública (SFP), que impulsa diversas líneas de reforma administrativa que contribuyen a crear un entorno más favorable. También a partir de 2001 la Secretaría de Desarrollo Social (SEDESOL) inicia un proceso de

institucionalización de la evaluación en el ámbito de las políticas sociales que culminará en 2005 con la creación de CONEVAL. La institución tiene la atribución de evaluar – y coordinar la evaluación- de programas y políticas en el ámbito social en todas las secretarías del gobierno federal. La conexión entre los componentes de monitoreo y evaluación se realizará a partir de la Ley de Presupuesto y Responsabilidad Hacendaria promulgada en 2006 y reformada en 2007, que crea el Sistema de Evaluación para el Desempeño (SED). El SED establece que la coordinación del monitoreo y la evaluación de programas federales estará a cargo del CONEVAL, la SHCP y la SFP. Esa coordinación, junto a las pautas de monitoreo y evaluación, se regula en los Lineamientos Generales para la Evaluación de Programas Federales publicados en marzo de 2007. Esos Lineamientos establecen acciones para la mejora del diseño de las políticas, fijando la matriz del marco lógico como método de referencia para asegurar la lógica interna de los programas. Además, se impulsa un Programa Anual de Evaluación (PAE) que incluye diferentes instrumentos de evaluación –de consistencia y resultados, de impacto, de procesos...- dejando margen a los diferentes organismos públicos para que realicen evaluaciones por su cuenta. También se introducen en esta etapa mecanismos para potenciar el uso de la información generada a través del sistema, que se publica regularmente en el web de CONEVAL. Se identifican como retos pendientes la mejora de la planeación y del diseño de las políticas, una mayor integración de los actores impulsores del sistema, la mejora de las condiciones técnicas de evaluación, así como del uso efectivo de los resultados en los procesos decisionales.

2.1. La evaluación en Colombia

El modelo actual de institucionalización de la evaluación en Colombia arranca del proceso de cambio político que culminó en la reforma constitucional de 1991. Sin embargo, es importante señalar algunos antecedentes de interés.

El modelo de planeación nacional instaurado en los años 50 empezó a mostrar signos de agotamiento y una necesidad urgente de reformulación en la década de los 80, a raíz del incremento de los presupuestos de inversión y su escaso impacto en el desarrollo económico y social del país. Con esta preocupación, a partir de 1990 se pusieron en marcha los convenios de gestión como mecanismos de validación de resultados para las entidades beneficiarias de créditos de organismos multilaterales de los que era garante el Gobierno.

Por otra parte, también a finales de los 80, el Departamento Nacional de Planeación (DNP), un organismo creado en 1958, con rango ministerial, pone en marcha dos iniciativas, antecedentes del sistema de monitoreo y evaluación actual (DNP, 2010):

- El Sistema de Evaluación y Seguimiento de Proyectos (SISEP), bajo dependencia de la Dirección del DNP y posteriormente integrado a la unidad de Inversiones Públicas, hoy Dirección de Inversiones y Finanzas Públicas.

- El Banco de Proyectos, creado por la Ley 38 de 1989, con la función de hacer “la evaluación de todos los proyectos de inversión financiados con recursos del gobierno central a fin de que cumplan con criterios básicos de rentabilidad económica y viabilidad técnica.”

En 1990-1991 el país atraviesa un proceso de cambio político que culminó en ese último año con una reforma de la antigua Constitución de 1886. La nueva Constitución, entre otros aspectos, suprime restricciones existentes hasta entonces a la competencia abierta entre distintas fuerzas políticas y establece, junto a los mecanismos clásicos electorales, dispositivos de democracia participativa como el cabildo abierto (o reunión pública de los órganos de gobierno local), la iniciativa legislativa o la posibilidad de revocación del mandato de los electos por parte de los ciudadanos.

Pero además, la Constitución de 1991 amplía notablemente la relación de derechos de los ciudadanos, sumando los sociales y económicos y también los llamados de tercera generación o derechos colectivos. Otras modificaciones que contribuyeron a crear un entorno más propicio para la evaluación fueron la despolitización de la Contraloría General de la Nación y la disminución del poder del Presidente a favor del Legislativo (Alcántara, 2013).

La Constitución establece también dos principios rectores de la planeación nacional: el primero es el carácter participativo del proceso de formulación de los planes de desarrollo; el segundo es el principio de evaluación continua de los planes y de la gestión pública en términos de resultados. (Ocampo, 1996, citado por Villarreal, 2007).

Primera etapa: 1994-2002. La Ley del Plan Nacional de Desarrollo y el lanzamiento de SINERGIA

En sus artículos 343 y 344, la Constitución Política de Colombia de 1991 fija el marco para diseñar y organizar los sistemas de evaluación de gestión y resultados de la administración pública⁵⁹. Es a partir de esta promulgación que la evaluación se introduce como instrumento de gestión del sector público, dejando de ser un ejercicio voluntario o aislado de algunas entidades (DNP, 2010).

Dentro de la serie de normativas jurídicas que desarrollaron y reglamentaron el proceso de seguimiento y evaluación en Colombia, destacan el Decreto 2167 de 1992; la Ley 42 de 1993 y la Ley 152 de 1994.

59. Artículo 343: “La entidad nacional de planeación que señale la ley, tendrá a su cargo el diseño y la organización de los sistemas de evaluación de gestión y resultados de la administración pública, tanto en lo relacionado con políticas como con proyectos de inversión, en las condiciones que ella determine.”

Artículo 344: “Los organismos departamentales de planeación harán la evaluación de gestión y resultados sobre los planes y programas de desarrollo e inversión de los departamentos y municipios, y participarán en la preparación de los presupuestos de estos últimos en los términos que señale la ley. En todo caso el organismo nacional de planeación, de manera selectiva, podrá ejercer dicha evaluación sobre cualquier entidad territorial.”

El Decreto 2167 de 1992 reestructura el DNP y le atribuye, entre otras, las funciones siguientes:

- elaborar las metodologías para la preparación y evaluación de los programas y subprogramas del Plan Nacional de Desarrollo,
- diseñar y organizar sistemas de evaluación de gestión y de resultados de la administración, tanto en lo relacionado con políticas como con proyectos de inversión, y señalar los responsables, términos y condiciones para realizar la evaluación, y
- organizar y poner en funcionamiento un sistema de evaluación retrospectiva del Plan Nacional de Desarrollo y de los planes de las entidades territoriales.

Por su parte, la Ley 42 de 1993, hacía referencia al control fiscal financiero y la manera como lo ejercería la Contraloría General de la república. De manera complementaria, en el mismo año se aprobó la Ley 87 de 1993, por la cual se establecieron normas sobre el ejercicio del control interno en las entidades del Estado.

En 1994, bajo la Presidencia de Ernesto Samper, se aprueba la Ley Orgánica del Plan de Desarrollo (Ley 152/94), siendo la primera vez que ese instrumento de planeación, presente en Colombia desde hacía más de tres décadas, adopta un rango de ley. De acuerdo con la Constitución, la norma otorga al Departamento Nacional de Planeación (DNP) la función de “diseñar y organizar los sistemas de evaluación de gestión y de resultados de la administración, tanto en lo relacionado con políticas como con proyectos de inversión, y señalar los responsables, términos y condiciones para realizar la evaluación. Dichos sistemas tendrán en cuenta el cumplimiento de las metas, la cobertura y calidad de los servicios y los costos unitarios, y establecerán los procedimientos y obligaciones para el suministro de la información por parte de las entidades.” (Art. 29).

También en 1994, el Consejo Nacional de Política Económica y Social (Conpes)⁶⁰ aprueba el Documento Conpes 2688 *La evaluación de resultados en el sector público a nivel nacional*, y dicta una resolución ese mismo año por la que se instaura el Sistema Nacional de Evaluación de resultados de la Gestión Pública (SINERGIA). Se concibe como un sistema nacional que integra diferentes actores –entidades gubernamentales, entidades de control del ejecutivo, la sociedad civil, los medios de comunicación, la academia, etc.- en actividades de monitoreo y evaluación.

Posteriormente, en 1995 la Consejería para la modernización del Estado y la División Especial de Evaluación (DEE), adscrita al DNP, elaboraron el documento Conpes 2790

60. Órgano creado en 1958, es la máxima autoridad nacional de planeación y actúa como organismo asesor del Gobierno en todos los aspectos relacionados con el desarrollo económico y social del país. Con esa finalidad, coordina y orienta a los organismos encargados de la dirección económica y social en el Gobierno, a través del estudio y aprobación de documentos sobre el desarrollo de políticas generales que son presentados en sesión. El DNP actúa como Secretaría Ejecutiva del Conpes.

Gestión pública orientada a resultados, en el que se destaca la autoevaluación por parte de las entidades a cargo de los programas públicos como concepto clave en SINERGIA. En línea con esa perspectiva, se define el proceso de elaboración de planes indicativos cuatrienales por parte de las entidades de la Administración Pública nacional, que se conciben como herramientas estratégicas para fijar objetivos y para realizar el monitoreo y la evaluación del grado de progreso alcanzado en relación a las metas fijadas en el Plan Nacional de Desarrollo⁶¹. En Septiembre de 1995 se puso en marcha una experiencia piloto de implantación de planes indicativos en 12 entidades de nivel nacional (DNP, 2010).

Derivado de lo anterior, se conformaron equipos de trabajo interinstitucionales denominados “Unidades de Gestión”, creados con el fin de promover la transversalidad del sistema y de concertar objetivos de los diferentes programas del PND, asegurando la coordinación por parte de los organismos involucrados, así como un seguimiento conjunto de su cumplimiento. Además, se buscó la participación del Ministerio de Finanzas para asegurar la correspondencia con la planificación presupuestaria. Sin embargo, si bien esta estrategia otorgó a los miembros de las unidades participantes la posibilidad de tener una visión más completa de su trabajo y contribuyó a resolver algunos problemas de coordinación interinstitucional, finalmente la iniciativa se interrumpió debido a su escasa relación con la asignación presupuestal y a la dificultad de lograr acuerdos sobre las competencias y responsabilidades de cada entidad (DNP, 2010).

El primer documento Conpes de seguimiento del PND, en 1997⁶², propuso un nuevo mecanismo, los “Acuerdos de Eficiencia”, orientado a mejorar la coordinación entre las instancias de planeación, de gestión de los programas y de gestión presupuestaria. Con ellos se buscaba asegurar una mayor certidumbre sobre los recursos presupuestarios disponibles, fijar de forma más clara las prioridades de gasto y avanzar en el cumplimiento de los resultados previstos.

No obstante, y a pesar de que ya en 1997 todas las entidades públicas con responsabilidad en el cumplimiento de metas del PND (alrededor de 170) realizaron un esfuerzo de definición de objetivos e indicadores en el marco de SINERGIA, la información tanto relativa a la planificación como al monitoreo presentaba carencias en cuanto a su calidad y factibilidad, además de persistir los problemas de conexión entre los planes, los mecanismos de monitoreo y evaluación y la gestión del presupuesto (DNP, 2010).

Al final de la década, el proceso empieza a verse afectado por las severas restricciones presupuestarias que se presentaron con la crisis fiscal de finales de siglo, dificultades a las que se suma el peso de la seguridad y la defensa nacional en la agenda política en

61. Los planes indicativos debían contener los objetivos del PND vinculados con la entidad, los objetivos prioritarios de la entidad para lograrlos, los indicadores de eficiencia y eficacia en términos de recursos presupuestarios y financieros, etc.

62. Documento Conpes 2917. *El Salto Social. Informe de avance del Plan Nacional de Desarrollo*. 1997.

ese mismo período. Así, entre 1998 y 2000 el Sistema estuvo en su punto más bajo de actividad, respaldo político y estructura institucional, por lo que el proceso de implementación sufrió un estancamiento.

Segunda etapa (2002-actualidad): ajustes y consolidación del sistema

La cultura de la gestión por resultados entra con fuerza en la Administración Colombiana en el primer mandato del Presidente Uribe, entre 2002 y 2006. Además del objetivo prioritario de recuperar la seguridad en el país, la Administración Uribe incorpora una agenda de reformas que gira en torno a tres pilares (DNP, 2010): una nueva cultura de gestión de lo público, una profundización de la descentralización y el fortalecimiento de la participación ciudadana.

En lo que respecta a la gestión por resultados, se aborda un proceso de reforma de SINERGIA, orientado a superar debilidades identificadas en el período previo, relativas a la conexión del sistema con la gestión del presupuesto, pero también al marco conceptual de monitoreo y evaluación, a la calidad y fiabilidad de la información obtenida y a la periodicidad de su producción, de forma que se asegurase una conexión óptima con los procesos de toma de decisiones.

Así las cosas, en el marco del Programa de reforma de la Administración Pública (PrAP), establecido a partir de los objetivos estratégicos marcados por la Directiva presidencial no. 10 de 2002, SINERGIA inició un proceso de renovación que buscó fortalecer el sistema en torno a tres grandes objetivos: 1) contribuir a mejorar la eficacia y el impacto de las políticas, programas e instituciones, 2) incrementar la eficiencia y transparencia en la programación y asignación de los recursos y 3) estimular la transparencia en la gestión pública, incorporando de manera activa el control de la ciudadanía.

En concreto, el sistema desarrolló los componentes siguientes:

a) El componente de monitoreo o seguimiento

Con este componente, a partir de 2002 Sinergia empezó a producir información de corto plazo sobre los resultados de las políticas y programas reflejados en el PND, a través de una plataforma tecnológica denominada SIGOB (Sistema de Gestión y Seguimiento a Metas de Gobierno).

El diseño original del SIGOB corresponde al Programa de Naciones Unidas para el Desarrollo (PNUD), en el marco del proyecto regional “Gestión para la Gobernabilidad”. Consiste en un sistema vía internet, que recopila información de los diferentes ministerios y entidades administrativas, y que además contempla la transparencia de la gestión a través de la publicación de los resultados y metas. Ese sistema estuvo activo entre 2003 y 2010, siendo reemplazado en 2011 por un nuevo sistema denominado

SISMEG (Sistema de Seguimiento a Metas de Gobierno), coincidiendo con la sucesión de Álvaro Uribe por el Presidente Juan Manuel Santos tras las elecciones de 2010.

El SIGOB persigue una integración de la información de metas y resultados de las instituciones principales de la rama ejecutiva a lo largo de un cuatrienio. De manera específica los objetivos del SIGOB fueron:

- Servir como herramienta de información y control de la gestión de las entidades por parte del Presidente.
- Brindar a la ciudadanía y a todas las instancias y ramas del poder público un acceso simultáneo y transparente sobre los resultados del Gobierno.
- Concentrar la información básica para el análisis ex-post de las principales políticas públicas, en aras de fortalecer los ejercicios prospectivos y de planeación futuros.
- Facilitar el control de la gestión de las entidades responsables del cumplimiento de las metas del Plan Nacional de Desarrollo y de los compromisos del Gobierno.

A mediados de la década de 2000, el sistema contenía cerca de 500 indicadores relacionados con 320 objetivos del PND. Para cada indicador, el sistema registraba el objetivo, las estrategias en curso para su consecución, información de la línea de base, objetivos operativos anuales, resultados en base a esos objetivos y recursos asignados y ejecutados en cada caso. La información se encontraba desagregada por región y también por algunas de las principales ciudades. De forma adicional, para reforzar la rendición de cuentas, en caso de objetivos no alcanzados, el gestor responsable debía ofrecer explicaciones al respecto, que eran incluidas en el sistema de información. Todos los gestores a cargo de programas quedaban identificados en el sistema, señalando la entidad en la que prestan servicios y sus datos de contacto (Mackay, 2007).

Por otra parte, en cuanto a áreas de mejora, si bien el sistema fue utilizado profusamente desde la Presidencia de la República para realizar el seguimiento del programa de Gobierno, la información presentaba problemas de calidad, fiabilidad y parcialidad, además de no obedecer a estándares homogéneos. Parte del problema provenía del hecho de que la información era introducida directamente en el sistema por parte de las diferentes entidades de la Administración, sin que existiese ningún sistema de auditoría o control de calidad de la misma.

El SIGOB, como decíamos, fue reemplazado por el SIMEG en 2011. Se trata también de un sistema de monitoreo de la planeación gubernamental, pero presenta algunas diferencias con relación a su antecesor. En primer lugar el SISMEG es una plataforma más avanzada tecnológicamente que incluye, aparte del sistema de seguimiento, dos productos disponibles a la consulta de la ciudadanía: la Bitácora Económica y el Comparativo Internacional⁶³. El sistema emite, entre otros, informes periódicos de balance de

63. <https://sinergia.dnp.gov.co/portaldnp/>

resultados, nacional y territorial, y un informe también anual para el Congreso de la República. Además de los objetivos de seguimiento y control de gestión internos, y de rendición de cuentas, el SISMEG se propone también mejorar la vinculación de la planeación con la asignación de recursos presupuestarios.

A lo largo de todo el período, el DNP, en coordinación con la Presidencia, ha implementado acciones para mejorar la calidad y actualización de la información contenida en los sistemas de seguimiento descritos, así como para incentivar su uso. Esas acciones han incluido formación a los organismos públicos en temas de seguimiento, asesoramiento directo, elaboración de manuales y guías, y también, finalmente, controles a las entidades en ejercicios de rendición de cuentas y controles de gestión, elaborar documentos de seguimiento a resultados e iniciar la articulación del Sistema de Seguimiento a metas de Gobierno con otros sistemas de información, entre otros.

b) El componente de evaluación

También a partir de 2002 se empieza a desarrollar el componente de evaluaciones de SINERGIA, con la finalidad de mejorar la efectividad, el diseño y la ejecución de políticas y programas incluidos en el PND.

Los estudios a realizar se reúnen en agendas de evaluaciones de vigencia anual. La construcción de la agenda es en primera instancia un ejercicio técnico entre las direcciones técnicas del Departamento Nacional de Planeación y las diferentes entidades del Gobierno, en el cual se identifican los temas prioritarios de acuerdo a los objetivos estratégicos del PND y los compromisos de evaluación consignados en los programas financiados por organismos multilaterales. El proceso es liderado por la Dirección de Seguimiento y Evaluación de Políticas Públicas (DEPP) del DNP.

A diferencia de CONEVAL (México), la DEPP no goza de autonomía técnica, de gestión y financiera, sino que, como la DIPRES (Chile) es un órgano administrativo, que forma parte a todos los efectos de la estructura de la DNP. Aunque no goza de la capacidad de *enforcement* del organismo chileno, directamente vinculado a la gestión del presupuesto.

En 2007, con el objetivo de dotar de sostenibilidad y continuidad a esta forma de funcionamiento se creó a través del artículo 132 de la Ley 1151 de 2007 (Ley del Plan Nacional de Desarrollo 2006-2010) el Comité Intersectorial de Evaluación y Gestión por resultados (CIE), con el fin de definir la agenda de evaluación plurianual, con base en las propuestas de las entidades. Así mismo, se definieron los criterios para priorizar los programas a evaluar y las metodologías a utilizar.

Una vez fijada la agenda, las evaluaciones se diseñan normalmente desde la DEPP o directamente por las entidades responsables de los programas, aunque el porcentaje de evaluaciones llevadas a cabo con participación de la DEPP respecto al total no ha

dejado de aumentar, llegando en 2010 a casi el 90%⁶⁴. La ejecución se realiza normalmente por parte de expertos académicos o consultores, a los que se adjudica al trabajo a través de licitaciones competitivas. Los manuales y directrices internas sobre cómo llevar a cabo evaluaciones enfatizan la importancia de involucrar a las partes interesadas en la evaluación, desde la fase de diseño, incluyendo el proceso de implementación y la difusión y debate de resultados.

El sistema de evaluaciones distingue diferentes tipos de evaluación, entre los cuales destacan las evaluaciones operativas y las evaluaciones institucionales, por un lado, centradas en aspectos de funcionamiento y gestión de los programas, así como de su contexto institucional; las evaluaciones de resultados, y las evaluaciones de impacto, siendo éstas últimas las más exigentes metodológicamente y comportan construcción de escenarios contrafactuales. También, evaluaciones ejecutivas, que son revisiones rápidas de un programa en funcionamiento para verificar la calidad de su diseño, cómo está funcionando, qué grado de cobertura está logrando, etc. Entre 2006 y 2011 se han llevado a cabo 19 evaluaciones ejecutivas, 22 de impacto, 7 institucionales, 5 de operaciones y 8 de resultados (61 en total) (DINP, 2010).

La distribución sectorial de las evaluaciones ha ido aumentando con el tiempo: mientras en 2004 el 44% de las evaluaciones estaban referidas a programas de protección social y el resto al sector de ambiente, vivienda y desarrollo territorial (AVDT), en 2010 se sumaban también la agricultura, la educación y el transporte y las infraestructuras, entre otros (DNP, 2010). En cuanto al volumen, hasta 2009, la agenda de evaluaciones liderada por la DEPP había desarrollado más de 60, invirtiendo cerca de 15 millones de dólares. El número de evaluaciones finalizadas en 2008 con respecto a 2004 presentó un incremento del 620%. En los últimos años, el número de evaluaciones anuales contemplado en las agendas ha oscilado alrededor de las 30⁶⁵.

Desde 2011 se está poniendo en marcha una nueva estrategia denominada Sistema Nacional de Evaluaciones (SISDEVAL). El objetivo es fortalecer metodológica y procedimentalmente las evaluaciones –a través de un componente de “evaluaciones efectivas” orientado a mejorar la capacidad técnica de los actores públicos y privados que intervienen en el proceso de evaluación-, así como inventariar y difundir públicamente las evaluaciones realizadas en todos los niveles de gobierno –a través del “radar de evaluaciones”–.

Cabe mencionar también, la creación, en 2009, de la Red de Monitoreo y Evaluación de Política Pública en Colombia⁶⁶, con el objetivo de establecer lazos entre actores vinculados

64. Ver DNP (2010). Gestión Pública por resultados en Colombia. SINERGIA. Presentación realizada en Zacatecas, el 27 de Abril de 2010. Disponible en <http://bit.ly/17nl7Ts>

65. Ver Agenda de Evaluaciones 2013. DNP, DEPP, en <http://bit.ly/1B2Pz2h>, y la relativa a 2014, en <http://bit.ly/1MLZRbX>

66. <http://redcolme.ning.com/>

al diseño, desarrollo y evaluación de políticas públicas en Colombia, especialmente en el ámbito local.

Valoración general y principales retos pendientes

Se identifican los principales elementos siguientes:

- El uso de la evaluación en los procesos de toma de decisiones sobre las políticas y programas por parte de los organismos a cargo de los mismos, así como en el proceso de gestión del presupuesto, presenta todavía márgenes de mejora importantes, aunque han hecho algunos avances al respecto, como sistematizar las recomendaciones en forma de planes de acción de mejora. Falta, no obstante, avanzar en el seguimiento de esas actuaciones.
- La vinculación directa de la DEPP con el Ejecutivo es percibida a veces desde instancias sociales y medios de comunicación como una fuente de parcialidad, y en general de falta de independencia del sistema de monitoreo y evaluación.
- La calidad de la información abocada al sistema por parte de los organismos públicos presenta todavía problemas de calidad en algunos casos. Hay también dificultades de disponibilidad y de calidad de datos en algunos sectores de intervención pública.
- Finalmente, la financiación de la agencia de evaluaciones depende todavía en gran medida de donantes multilaterales, lo que amenaza su sostenibilidad económica a medio y largo plazo.

El sistema de evaluación en Colombia

Valoración general

El impulso de la evaluación en Colombia parte del proceso de cambio político que culmina en la nueva Constitución de 1991, que asigna de forma explícita responsabilidades de evaluación a la entidad responsable de la planeación nacional, el Departamento Nacional de Planeación (DNP). En 1994 se aprueba la Ley Orgánica del Plan de Desarrollo y se pone en marcha el Sistema Nacional de Evaluación de resultados de la Gestión Pública (SINERGIA), que incorpora un componente de monitoreo y otro de evaluación. Sin embargo, el despliegue del SINERGIA entra en una fase de estancamiento a finales de los 90, de la que no saldrá hasta 2002, cuando bajo el mandato del Presidente Uribe se impulsa con fuerza la mejora de la gestión pública orientada a resultados. El sistema de evaluación y monitoreo de Colombia se ha diseñado, a diferencia de Chile, con capacidades ejecutivas limitadas. Sin embargo, desde el DNP, la Dirección de Seguimiento y Evaluación de Políticas Públicas (DEPP) ha ido legitimando y afianzando su papel de organismo de referencia en evaluación, capaz de asesorar y de proveer apoyo metodológico a las diferentes agencias del gobierno en sus actividades de evaluación (confección de términos de referencia, contratación y gestión de evaluaciones, asesoramiento metodológico en

evaluaciones de impacto, etc.). La DEPP ha seguido una estrategia de aproximación gradual, basada en la voluntariedad y en la concienciación y convencimiento progresivos de los organismos gubernamentales sobre el valor que aporta la evaluación para la mejora del diseño, del funcionamiento y de la efectividad de los programas. La difusión de los resultados del sistema de monitoreo y evaluación es sistemática, y la utilización de la información derivada de SINERGIA parece elevada por parte de la Presidencia y del Congreso. Pero su utilización efectiva en la mejora de los programas y en la asignación de recursos presupuestarios presenta márgenes de mejora, así como la calidad de la información abocada al sistema en algunos casos. Por último, la financiación en gran parte externa del sistema plantea un reto de sostenibilidad a medio plazo.

Tabla 3. Visión panorámica de los sistemas de evaluación descritos

Parámetros característicos de los sistemas de evaluación	Alemania	Bélgica	Chile	Colombia	España	Francia	Italia	México	Países Bajos	Reino Unido	Suecia
Marco conceptual											
– Práctica de evaluación de políticas públicas diferenciada de monitoreo.	•	?	•	•	•	•	•	•	•	•	•
– Evaluaciones de impacto con diseño experimental o cuasi-experimental.	•		•	•	•	•	?	•	•	•	•
– Orientación al control	•	•	•	•	•	•	•	•	•	•	•
– Orientación al aprendizaje y la mejora	•	•	•	•	•	•		•	•	•	•
Marco institucional - normas											
– Política explícita de evaluación de políticas públicas.			•	•		?		•	?		
– Referencia a la evaluación en normas transversales (presupuesto, planeación...).	?	?	•	•	•	•	?	•	•	•	?

Tabla 3. Visión panorámica de los sistemas de evaluación descritos (cont.)

Parámetros característicos de los sistemas de evaluación	Alemania	Bélgica	Chile	Colombia	España	Francia	Italia	México	Países Bajos	Reino Unido	Suecia
– Cláusulas de evaluación en legislación sectorial.	•	•	?	?	•	•	?	•	?	?	?
Marco institucional - organización											
– Monopolio			•								
– Pluralismo centralizado				•		•		•	•	•	
– Pluralismo	•										•
– Concurrencia fragmentada		•			•		•				
– Existencia de unidades de evaluación diferenciadas en organismos públicos			•	•	•	•	•	•	•	•	•
Práctica evaluadora – ejecución											
– Interna, centralizada					•	•		•		•	
– Interna, descentralizada		?			•	•	•		•	•	•
– Externa	•	?	•	•	•	•	•	•	•	•	•
Práctica evaluadora – calidad											
– Aplicación de estándares comunes			•	•				•		•	
– Supervisión experta	•		•	•		•		•		•	
– Disponibilidad y calidad de datos	•	?	?	?		?	?	?	•	•	•
Práctica evaluadora - continuidad											
– Agendas de evaluación	?		•	•	•	?		•	•	?	?
– Ejecución regular de evaluaciones	•		•	•		•		•	•	•	•
– Presupuesto específico	•	?	•	?	•	?	?	?	?	?	?

Tabla 3. Visión panorámica de los sistemas de evaluación descritos (cont.)

Parámetros característicos de los sistemas de evaluación	Alemania	Bélgica	Chile	Colombia	España	Francia	Italia	México	Países Bajos	Reino Unido	Suecia
Difusión pública											
– Difusión generalizada	•	?	•	•	•	•	?	•	•	•	•
– Difusión restringida		?					?				
– Sin difusión pública		?					?				
Utilización efectiva											
– En la mejora de los programas	•		•	•		•		•	•	•	•
– En el control parlamentario del ejecutivo	•		•	?		•		•	•	•	•
– En el control del ejecutivo por parte de la sociedad civil			?	?		?		?	?	•	•
– Instrumental, como baza para lograr consensos		•									
– Baja utilidad efectiva		•			•		•				
Evaluación del sistema											
– Existencia de organismos independientes que analizan regularmente el funcionamiento del sistema.									•	•	•
– Seguimiento interno, a cargo de los órganos responsables del sistema.			•	•				•	•		
– Sin seguimiento sistemático	•	•			•	•	•				

Tabla 3. Visión panorámica de los sistemas de evaluación descritos (cont.)

Parámetros característicos de los sistemas de evaluación	Alemania	Bélgica	Chile	Colombia	España	Francia	Italia	México	Países Bajos	Reino Unido	Suecia
• De la documentación consultada se desprende que el parámetro en cuestión se da en el país*											
De la documentación consultada se desprende que el parámetro en cuestión no parece darse en el país*											
? La documentación consultada no permite establecer si el parámetro en cuestión se da o no en el país*											

(*) El contenido de la tabla se basa en el análisis de fuentes secundarias seguido en la elaboración de este informe, las cuales no permiten capturar en todos los casos, ni con el mismo grado de certeza y detalle, información sobre los componentes de los sistemas de evaluación identificados en la literatura. Queda para más adelante, el contraste riguroso de las percepciones aquí recogidas, que debería realizarse partiendo de una revisión crítica de los componentes de los sistemas de evaluación identificados en la literatura, para establecer posteriormente, de forma razonada, variables e indicadores para cada uno de ellos, así como métodos de recogida de información que permitan medirlos adecuadamente, de forma homogénea y sistemática.

Fuente: elaboración propia.

V. Conclusiones y recomendaciones

El desarrollo exitoso de sistemas de evaluación comporta no tanto un cambio técnico o un cambio institucional. Implica, ante todo, un cambio cultural.

Ese **cambio cultural** debe comenzar por la base, es decir, por reivindicar los valores que deben guiar la acción política y que deben orientar la función de las administraciones públicas. Las élites políticas deben priorizar y estimular la reflexión y el debate sobre qué constituye el interés general y cuáles son las políticas públicas que mejor lo sirven. Un debate difícil, conflictivo, donde raramente habrá consenso y donde las reglas del juego democrático deberán determinar qué opciones van a gobernar la comunidad en cada momento. Pero infinitamente más necesario que la confrontación puramente táctica y cortoplacista que aleja de la política a los ciudadanos.

Ese cambio cultural no es sencillo, ya que va en contra de unas reglas del juego según las cuales la excelencia tiene menor repercusión que el fracaso, y éste, se castiga de tal modo que provoca aversión al riesgo y por tanto menos oportunidades de experimentación, de aprendizaje y de mejora:

“... within the political arena, the punishment for mistakes is severer than the reward for excellence.... Mistakes are more newsworthy and therefore get much more attention. This leads to slightly neurotic tensions in public organizations to avoid mistakes as much as possible. It is very well possible to have a successful career in the public sector by avoiding risks, whereas one mistake can kill a career. It also strengthens the tendency towards bureaucracy; rules provide a shield against criticism.” (OECD, 2009)

En cualquier caso, esa **voluntad política de reflexión rigurosa sobre qué constituye el interés general y cuáles son las políticas que mejor lo sirven**, y de hacer lo posible por avanzar al respecto, es el motor fundamental de arranque y desarrollo de la evaluación. Sin esa voluntad política, nacida del **convencimiento personal** sobre la necesidad de mejorar la acción de gobierno para perseguir el bien público, y **sobre el valor intrínseco que aporta la evaluación a esa tarea**, la mera introducción de normas, de organismos o de técnicas de evaluación en las administraciones difícilmente llegará a echar raíces, más allá de algunas demasiado débiles para evitar que el sistema

se derrumbe o debilite irremisiblemente ante cualquier cambio político o nueva moda de reforma de la Administración.

Para desarrollar con éxito un sistema de evaluación, esa voluntad de partida y ese convencimiento personal deben encontrarse en una multiplicidad de individuos y ser compartida por todos ellos. Y esos individuos no deben pertenecer sólo a las élites políticas. También deben encontrarse entre los profesionales del sector público que gestionan programas, recursos, que prestan servicios directamente a los ciudadanos. Y entre los expertos académicos y entre las entidades de la sociedad civil, porque el desarrollo de la evaluación necesita la participación de todos ellos.

Generar una cultura política, social y administrativa fértil para la evaluación requerirá, por tanto, invertir tiempo y recursos en **sensibilizar**, en **explicar**, en **formar** a todos esos actores sobre qué es la evaluación de políticas públicas y por qué es interesante adoptarla en ese ejercicio permanente de definición y de prosecución del bien común. No estamos refiriéndonos todavía a una formación técnica sobre metodología evaluadora, sino a acciones de difusión, a seminarios, sesiones de trabajo, talleres introductorios, conferencias, reuniones, etc., donde se explique de forma sencilla qué es la evaluación, qué puede aportar, qué implica; donde se presenten de forma atractiva comprensible resultados de evaluaciones y se sometan a debate entre expertos, gestores, beneficiarios y ciudadanos.

Esas actuaciones de sensibilización y de formación introductoria contribuyen, además, a la adopción de una base epistemológica común. Es decir, un **entendimiento compartido sobre qué es la evaluación de políticas públicas**, en qué se diferencia de otras técnicas afines descritas en apartados previos, cuál es su aportación, qué puede esperarse de ella.

Sobre la base de esa voluntad política y de ese universo cognitivo compartido, pueden erigirse con cierta solidez el resto de piezas del sistema vistas en apartados previos... aunque siempre contando con ciertos factores del contexto que acompañen adecuadamente el proceso, tal como vimos en apartados previos.

Uno fundamental reside en la **academia**, y en general en los focos de conocimiento experto en problemáticas sociales, en políticas públicas, en métodos de investigación y en evaluación de políticas públicas. Para que la evaluación de políticas públicas se desarrolle, esos focos de conocimiento experto deben existir y deben alcanzar una masa crítica adecuada para poder garantizar una cobertura óptima respecto a la necesidad de evaluaciones externas.

Pero su mera **existencia, masa crítica y calidad** no es suficiente. A menudo ese conocimiento experto adolece de una excesiva compartimentación en diferentes disciplinas académicas, escasamente conectadas entre sí. La evaluación requiere, en cambio,

como vimos, hibridación, perspectivas y **equipos multidisciplinares, métodos mixtos** de investigación. Por otra parte, la conexión entre ese conocimiento experto y la administración no es evidente. La transferencia de conocimiento desde la academia hacia la administración no está suficientemente articulada ni incentivada.

Las administraciones olvidan a menudo incluir en sus convocatorias de ayudas a la investigación líneas específicas orientadas a responder preguntas sobre la idoneidad de intervenciones públicas, a profundizar en el conocimiento sobre determinadas problemáticas sociales que ocupan los primeros puestos en agendas políticas. Por otro lado, la carrera académica, basada en una producción científica de calidad, no parece contar con suficientes incentivos para la transferencia de conocimiento al sector público. Los tiempos, las preguntas, los contenidos y los formatos adecuados a los estándares de la producción científica, por una parte, y a los procesos de toma de decisiones sobre las políticas públicas, por la otra, son a menudo del todo distintos. Habrá que trabajar también en ese ámbito, para establecer **mecanismos que favorezcan la transferencia de conocimiento, y en general una relación más fluida entre academia y administración**, entre conocimiento experto y procesos de toma de decisiones sobre las políticas públicas.

Pero además, hay que **fomentar la investigación evaluativa más allá de los proyectos puntuales, procurando el avance del conocimiento de las problemáticas sociales, así como la generación de modelos de análisis e instrumentos para su evaluación**. Es pues necesario estimular, organizar y dotar económicamente esas líneas de investigación y transferencia de *know-how* al sector público.

En cuanto a las **fórmulas de institucionalización**, hemos visto que la práctica de la evaluación en entornos gubernamentales puede adoptar múltiples formas. De hecho, puede desarrollarse sin que existan forzosamente normas y organizaciones transversales y específicas sobre evaluación, en entornos donde factores culturales promueven su práctica regular sin necesidad de ese marco institucional (p. ej., en Alemania). Puede darse también una combinación de factores culturales y políticos favorables, vinculados a su vez a una cierta regulación *ad hoc* de la evaluación y al apoyo y difusión continua de su práctica a través de incentivos y capacitación (p.ej., en el Gobierno Federal de Estados Unidos, que queda fuera del ámbito de este estudio). También, contextos culturalmente favorables donde la evaluación ha quedado durante un tiempo en un segundo plano respecto a prácticas de monitoreo orientado al control (p. ej., en el Reino Unido). Así mismo, encontramos entornos cultural y políticamente aún poco propicios donde primeros pasos orientados a establecer normas e instituciones formales no han permitido, por el momento, un avance rápido (p. ej., en Bélgica, Italia y España).

No obstante, **un cierto grado de institucionalización formal parece importante** para que la evaluación pueda jugar un papel en entornos gubernamentales, y en particular:

- Para otorgarle visibilidad y transparencia.
- Para protegerla, al menos formalmente, contra presiones del sistema político y burocrático que puedan desnaturalizarla (proclamando, por ejemplo, su independencia o la obligación de difundir públicamente sus resultados).
- Para facilitar, a través del establecimiento de un diseño y de unas pautas de funcionamiento explícito, el aprendizaje sobre su desarrollo, formas de implantación y conexión con procesos de toma de decisiones y rendición de cuentas.

Y también, por supuesto, para **fomentar una práctica regular previsible**, que siga **estándares adecuados de calidad** y que promueva su **utilización** efectiva.

Veamos con más detalle los diferentes elementos que componen ese proceso de institucionalización, y que incluyen normas, organización y práctica evaluadora.

En primer lugar, en lo que se refiere a la **regulación** o existencia de normas específicas sobre evaluación, no parece necesaria una regulación prolija. Ésta suele recogerse, por un lado, de forma transversal, a través de disposiciones presentes en normas con rango de ley relativas a las finanzas públicas o a la planeación nacional, y posteriormente es objeto de directrices internas a través de órdenes, circulares u otros documentos que dictan la política de evaluación. En esas normas transversales no es tan importante regular los detalles de la práctica evaluadora –que, como hemos visto, a menudo se ejerce de forma descentralizada por los diferentes departamentos y agencias del gobierno– como incidir en algunos elementos clave: sus objetivos, criterios y estándares de calidad, los recursos que se le asignan, el uso que debe hacerse de sus resultados y aspectos relativos a su difusión pública.

Vimos también entre los instrumentos de institucionalización normativa el recurso, en algunos países, a cláusulas legislativas que imponen obligaciones de evaluación. Se trata de disposiciones que entran raramente en detalles y se sitúan a un nivel general que abre la puerta a múltiples perspectivas de implementación.

Queda claro que no existe una pauta única de institucionalización normativa, aunque la experiencia aconseja evitar aproximaciones excesivamente legalistas, que impongan obligaciones de evaluación poco adaptadas a organizaciones escasamente receptivas o todavía no capacitadas técnicamente para poderlas asumir, ya que conllevan un elevado **riesgo de burocratización** de la evaluación. Hemos visto a lo largo del informe la diversidad de los procesos de institucionalización de la evaluación, sus dinámicas, fuertemente idiosincrásicas, su necesidad de implantación paulatina, gradual. En ese contexto, nada más contraproducente que confundir el impulso político del proceso con la simple imposición de nuevas obligaciones normativas a unas administraciones que no cuenten con la capacidad, con la motivación ni los recursos necesarios para asumirlas.

En cuanto a los **aspectos organizativos**, veíamos en apartados previos del informe diversas configuraciones organizativas posibles, que parecen determinadas en gran parte por las características del sistema político en el que se inscriben:

- Una configuración monopolística, donde una única organización ostenta el rol dominante en el impulso y en la práctica de la evaluación con una aproximación fuertemente normativa y “de abajo arriba”. Es una configuración propia de países con un elevado nivel de centralización y de etapas incipientes de desarrollo de la evaluación. Lo hemos visto en el caso del Reino Unido en relación no tanto al sistema de evaluación sino al de *Performance Measurement*, y en sus primeras etapas de implantación. También en Chile, que siguió al principio claramente el modelo anglosajón, aunque actualmente el sistema se ha hecho más flexible.
- Una situación de “pluralismo centralizado”, en el cual un organismo ocupa una plaza preponderante, determinando el comportamiento del resto pero combinando normas e incentivos, de forma que se dé una adhesión más o menos voluntaria y gradual al sistema. Es una configuración que vimos en el caso de los Países Bajos, del Reino Unido –aunque en este caso vinculado a la evaluación–, en Colombia, en Francia y, en menor medida, en México, donde existen diversos focos de impulso de la evaluación que actúan, en principio, de forma coordinada.
- Un escenario pluralista, propio de países fuertemente descentralizados como Alemania y Suecia.
- Un escenario de concurrencia fragmentada, propio de países con desarrollos incipientes y débiles de sus sistemas de evaluación, que aspiran a un modelo pluralista, o bien pluralista centralizado sin conseguirlo todavía plenamente. Los sistemas de evaluación de Italia, España y Bélgica parecen encajar en esta categoría.

En general, se ha puesto de manifiesto que la institucionalización de políticas públicas adopta múltiples formas, que se traducen en una diversidad de organismos, instancias o grupos más o menos complejos (institutos de investigación públicos y privados, administración central, regional, local, organismos autónomos, poder legislativo, poder ejecutivo, tribunales de cuentas, asociaciones de evaluación, consultoría privada, entidades filantrópicas, ONG, etc.), tanto a nivel central y regional o local en cada país, como también en el ámbito supranacional.

En el interior de la Administración, la **combinación de instancias de impulso, coordinación y supervisión** centralizada de la evaluación, que actúan también como **centros de recursos y apoyo metodológico, con una función evaluadora descentralizada**, adaptada de forma más o menos flexible a cada contexto organizativo, parece favorecer el avance de la evaluación.

Por otra parte, es preciso dotar al personal técnico de la administración que deba gestionar o utilizar la evaluación de **capacidades analíticas y de gestión del conocimiento**. La capacidad evaluadora no es fácil de improvisar. Transformar rápidamente

funcionarios de la administración con perfiles ajenos a la investigación en evaluadores expertos no es factible –sorprende ver con qué facilidad se ha asumido lo contrario en algunos de los casos descritos–, si tenemos en cuenta la sólida formación que éstos deben tener en métodos y técnicas de investigación social. No se trata tanto de formar dentro de la Administración evaluadores expertos de alto nivel –éstos se encontrarán mayoritariamente en la academia– sino impulsores y gestores de la evaluación, capaces de transformar las necesidades de conocimiento en preguntas abordables desde la evaluación, de actuar como clientes inteligentes de encargos externos, de hacer el seguimiento de la utilización de los informes y, si cabe, más importante, de elaborar planes de evaluación, velar por la consecución de óptimas condiciones internas de evaluabilidad y cultivar una cultura interna favorable. En cualquier caso, se trata de una función específica, que no debe asumirse como una responsabilidad añadida, superpuesta a otras funciones del puesto de trabajo, sin formación, sin recursos y sin ninguna base de legitimidad para ejercerla. Al contrario, hay que mimar esa función estratégica, encargada de la producción y gestión de conocimiento de calidad para la mejora de las políticas públicas. Hay que **definir las funciones organizativas relacionadas con el sistema de evaluación** y hay que **situarlas adecuadamente en los organigramas y en los catálogos de puestos de trabajo** de las administraciones públicas. Hay que definir sistemas adecuados de selección del personal, y de desarrollo profesional.

En cuanto a la **práctica evaluadora**, es importante asumir que **el progreso de la evaluación es gradual, irregular y asimétrico**. Por ello, parece adecuado fomentar **focos de evaluación de calidad** allá donde se da una combinación propicia de factores y procurando difundir posteriormente esas experiencias punteras entre las organizaciones del sector público, a través de **redes internas o interadministrativas**. Esas redes deberían de favorecer un “efecto contagio” positivo y un grado más elevado de implicación y de compromiso con la evaluación por parte de los profesionales del sector público. A lo largo del informe hemos visto a menudo ese tipo de “focos” de evaluación (como las evaluaciones de PROGRESA en México, o las relativas a políticas activas de empleo en Alemania, o de empleo juvenil en Francia, etc.), así como el establecimiento de esas redes en prácticamente todos los países estudiados.

Impulsar la práctica de la evaluación en las administraciones implica también comprometerse seriamente con la mejora de las condiciones actuales de evaluabilidad de las políticas, y en particular con la **disponibilidad y calidad de los datos contenidos en los registros administrativos**, con su **aprovechamiento estadístico**, y con la **facilidad de acceso** por parte de los profesionales de la evaluación. Las debilidades actuales en ese terreno plantean ahora retos a menudo insalvables para la realización de evaluaciones retrospectivas de impacto. Para avanzar hacia esas mejoras, es imprescindible la colaboración y la coordinación de estrategias y actuaciones entre las instancias responsables de impulsar la evaluación, los organismos a cargo de la estadística oficial, los órganos gubernamentales competentes en sistemas de información y TIC y

las instituciones responsables de la protección de datos personales. El impulso actual a políticas de gobierno abierto y de administración electrónica ofrece una valiosa oportunidad para desarrollar buenas bases de datos administrativas, que puede favorecer tanto a la evaluación como a la estadística oficial. Sin embargo, las sinergias entre esas iniciativas a menudo se descuidan.

Planificar de antemano las evaluaciones y abordarlas siempre que sea posible de forma prospectiva, junto con una actuación coordinada entre los organismos mencionados, comportaría, sin duda, mejoras ostensibles en las condiciones de evaluación a medio plazo. Con este fin es útil estimular la realización de **planes o agendas de evaluación, de carácter anual o plurianual**, una práctica que ya hemos visto en países como Chile, Colombia, México y los Países Bajos, y presente también en otros que quedan fuera del ámbito de este estudio (Estados Unidos, Canadá, Suiza). Esos planes de evaluación pueden realizarse por parte de instancias responsables de la evaluación en entornos gubernamentales centralizados, o bien por cada organismo o agencia gubernamental en entornos descentralizados.

La confección de esos planes o agendas de evaluación requiere la **priorización**, a efectos de evaluación, de un determinado número de políticas y programas sobre la base de una serie de criterios que habitualmente incluyen aspectos de relevancia política (relación con objetivos estratégicos del Gobierno, segmento poblacional afectado, etc.), de presupuesto (volumen de presupuesto asignado, dudas sobre la relación coste-beneficio / coste-efectividad, etc.) y de viabilidad de las evaluaciones (claridad de los objetivos y del diseño de la política, disponibilidad de información y datos, compromiso de los responsables políticos y técnicos del programa ante la evaluación, etc.). Junto a esos criterios principales, la confección de agendas de evaluación debe tener en cuenta otros relativos a características del programa que ayudan a determinar el enfoque de evaluación a emplear (si se trata de un programa nuevo o bien de una iniciativa ya madura; resultados de evaluaciones anteriores, etc.).

Esos planes o agendas de evaluación deberán incluir, por tanto, **qué programas van a evaluarse y qué tipo de evaluación va a aplicarse en cada caso** –de diseño, de implementación, de impacto retrospectiva o prospectiva, evaluación económica-. También deberían incluir la **programación en el tiempo** de esas evaluaciones y **asignar responsabilidades** –qué unidad va a hacerse cargo de impulsar y de gestionar la evaluación-, **recursos** –presupuesto- y forma de ejecución –interna o externa, y en este último caso tipo de licitación en función del presupuesto-.

Los planes de evaluación deberían recoger así mismo las **reglas de difusión pública** de las evaluaciones y los **mecanismos previstos para el seguimiento de la aplicación de los resultados y recomendaciones** que emanen de ellas –idealmente haciéndose eco de una regulación transversal existente al respecto-.

La existencia de políticas bien diseñadas, de planes o agendas de evaluación dotados de recursos económicos y técnicos adecuados, y la disposición de condiciones de evaluabilidad adecuadas no es tampoco suficiente para garantizar una práctica evaluadora regular y de calidad. Falta todavía un factor clave, que es la **demanda** de evaluaciones por parte de las instancias políticas y/o de los órganos gestores de los programas. No olvidemos que los órganos gestores y sus responsables tienen, en principio, pocos incentivos para someter a evaluación sus programas y rendir cuentas sobre su funcionamiento. Además, hemos visto que sistemas fuertemente centralizados, donde la decisión de qué, cómo y cuándo evaluar se impone de arriba abajo, implican problemas de apropiación de las evaluaciones, que comportan más tarde déficits en la **utilización de los resultados** para la mejora de los programas. Una evaluación de alta calidad técnica pero cuyos resultados no sirven para mejorar la política es una evaluación inútil. Para evitar ese problema, los sistemas de evaluación deben incluir **incentivos que estimulen la demanda y la apropiación** de la evaluación por parte de los órganos gestores de los programas. Y, en la medida de lo posible, deben ser incentivos positivos, que contribuyan a generar una cultura interna y unas condiciones técnicas favorables a la evaluación, además de facilitar el acceso a los recursos económicos necesarios. Un ejemplo al respecto pueden ser fondos concursales que ofrecen cofinanciación de proyectos de evaluación, de actividades de capacitación interna y de mejora de las condiciones de evaluabilidad. Ello sin perjuicio de incluir también otro tipo de incentivos, tanto positivos como negativos, vinculados en particular a la asignación de recursos presupuestarios.

Es también importante **implicar en el proceso evaluador a todos los actores relevantes**, tanto responsables políticos, como gestores del programa, otros organismos públicos o externos implicados en su diseño o implementación, beneficiarios, etc. Esa implicación cubre diversos objetivos: en primer lugar, contribuye a una mayor apropiación de la evaluación y, por tanto, puede facilitar la utilización posterior de los resultados en la toma de decisiones. Además, la participación de esos actores a lo largo del proceso constituye una ocasión excelente de sensibilización y de formación sobre qué es la evaluación, para qué sirve, cómo se desarrolla, etc.

En cuanto a la práctica de la evaluación, es también necesario adoptar mecanismos que permitan garantizar la **neutralidad y la calidad de las evaluaciones**. Ese control de calidad puede organizarse de formas diferentes, tal como hemos visto a lo largo del informe. Por ejemplo, a través de grupos o comités de expertos externos creados *ad hoc* para cada evaluación, que certifican la calidad de su diseño y ejecución, así como la idoneidad de los resultados y recomendaciones. O bien comités con esas mismas funciones y características, pero creados de forma más estable, para supervisar el conjunto de evaluaciones que se realizan, aunque su composición pueda adaptarse en cada caso. O, también, ese control de calidad puede ejercerse a través de organismos independientes de control y vigilancia del ejecutivo, que realizan regularmente metaevaluaciones de evaluaciones realizadas o encargadas por

aqué en determinados ámbitos de intervención. También, esa misma función y actividades pueden ser realizadas por organismos multilaterales, iniciativas internacionales de impulso de la evaluación o centros independientes de investigación. La difusión pública de las evaluaciones, así como el acceso a las bases de datos utilizadas en las mismas –con las garantías necesarias de protección de datos personales– puede facilitar un control de calidad desde múltiples instancias, incluyendo instancias sociales.

Dado que en este informe hemos abordado un enfoque sistémico de la evaluación, no debe bastarnos con asegurar el control de calidad de cada evaluación, sino que debemos aspirar también a contar con **mecanismos capaces de supervisar el funcionamiento global del sistema** y de introducir los ajustes y correcciones que sean necesarios en los diferentes elementos que lo componen, para procurar un funcionamiento adecuado. Hemos visto que en los países con mayor tradición evaluadora, los propios órganos de impulso de la evaluación en el ejecutivo o bien órganos superiores de control realizan un seguimiento periódico del progreso de la evaluación desde una perspectiva global, analizando cómo se lleva a cabo, los recursos que consume y, en particular, su incidencia en la mejora y en la transparencia de la acción de gobierno. Parece importante incluir en esa función de “supervisión sistémica”, la observación continuada de novedades en la configuración y funcionamiento de los sistemas de evaluación y en las políticas y estrategias que se despliegan en el entorno internacional para impulsarlos. Este es, como hemos visto, un ámbito nuevo, emergente en la literatura sobre evaluación y sobre gestión pública. Hacen falta líneas específicas de investigación al respecto, así como dar continuidad a las iniciativas de algunos organismos multilaterales, como el Banco Mundial o el BID, de describir y analizar el funcionamiento de dichos sistemas.

Finalmente, como decíamos en apartados de previos de este informe, entender qué es la evaluación, cuáles son sus activos y sus limitaciones, y apostar por ella, implica acometer un proceso de **transformación paulatina y profunda del modelo actual de administración pública**. Esa transformación debe permitir superar visiones estrictamente administrativistas –centradas en el necesario control de legalidad y seguimiento del procedimiento administrativo– o gerencialistas –enfocadas a la provisión de servicios de calidad y al control de la gestión–, para asegurar y proclamar la orientación finalista de todos los aspectos de la actuación administrativa. En otras palabras, debe asegurarse que las reformas se orientan en primera instancia a conseguir una actuación pública más efectiva a la hora de mitigar problemas de toda índole que reclaman la intervención de los poderes públicos. Para ello es necesaria una Administración capaz de apoyar y de gestionar la generación y la utilización continua de conocimiento de calidad sobre esos problemas y esas intervenciones. Si esa es la Administración que queremos, habrá que acometer en algún momento una revisión en profundidad de las políticas actuales de reforma administrativa, para asegurar que priorizan visiones y medidas orientadas a la consecución de ese objetivo.

Anexo I. Bibliografía

- ALCÁNTARA SÁEZ, M. (2013). *Sistemas políticos de América Latina*. Vol. I. América del Sur. Tecnos (4ª ed.).
- ARGILÉS, J.M. (2014). "Veinticinco años de evaluación en la política española de cooperación internacional para el desarrollo", en *Revista de Evaluación de Programas y Políticas Públicas*, No. 3, pp.19-52.
- ARENAS, A.; BERNER, H. (2010). *Presupuesto por Resultados y la Consolidación del Sistema de Evaluación y Control de Gestión del Gobierno Central*. Gobierno de Chile. Ministerio de Hacienda. Dirección de Presupuestos. Febrero 2010.
- BARBIER, J-C.; HAWKINS, P., (Eds.). (2012). *Evaluation Cultures. Sense Making in Complex Times*. Comparative Policy Evaluation, Vol. 19. Transaction Publishers. New Brunswick.
- HORBER-PAPAZIAN, K.; JACOT-DESCOMBES, C.; (2012), «Is Evaluation Culture Shaped by the Swiss Political System and Multiculturalism?», en BARBIER, J-C.; HAWKINS, P., (Eds.) (2012).
- SPEER, S. (2012), "Sectoral Evaluation Cultures: a Comparison of the Education and Labor Market Sectors in Germany", en BARBIER, J-C.; HAWKINS, P., (Eds.) (2012).
- STAME, N. (2012). "The Culture of Evaluation in Italy", en BARBIER, J-C.; HAWKINS, P., (Eds.) (2012)
- BASSANINI, F. (2009). *Twenty years of administrative reform in Italy*. <http://bit.ly/1aqphwY>
- BECK JØRGENSEN, T.; BOZEMAN, B. (2007). "Public Values. An Inventory". *Administration & Society*. Volume 39, No. 3, May 2007 354-381.
- BEMELMANS-VIDEC M.-L., ERIKSEN B., GOLENBERG N. (1994). "Facilitating organizational learning: human resource management and program evaluation", en LEEUW F., RIST R., SONNICHSEN R., (Eds.), *Can Government Learn? Comparative Perspective on Evaluation and Organizational Learning*, New Brunswick, Transaction Publishers, 1994
- BOHNI NIELSEN, S.; WINTHER, D. M. (2014). "A Nordic evaluation tradition? A look at the peer-reviewed evaluation literature". *Evaluation*. Vol. 20(3) 311–331.
- BOYLE, R.; LEMAIRE, D.; RIST, R. (1999). "Introduction: building evaluation capacity", *Building Effective Evaluation Capacity: Lessons from Practice*, Transaction Publishers. New Brunswick.
- BRYSON, J; CROSBY, B.C.; BLOOMBERG, L. (2014). "Public Value Governance: Moving beyond Traditional Public Administration and the New Public Management". *Public Administration Review*, Vol. 74, Iss. 4, pp. 445–456.
- BUSSMANN, W. (2008). "The Emergence of Evaluation in Switzerland". *Evaluation*. Vol 14(4): 499 –506.
- CASADO, D. (2009). *Guía práctica 2. Evaluación de necesidades sociales*. Colección Ivàlua de guías prácticas sobre evaluación de políticas públicas. Instituto Catalán de Evaluación de Políticas Públicas (Ivàlua).
- CHELIMSKY, E. (2008). "A Clash of Cultures: Improving the 'fit' Between Evaluative Independence and the Political Requirements of a Democratic Society". *American Journal of Evaluation*, núm. 29, p. 400-415.
- CHELIMSKY, E. (2009). "Integrating Evaluation Units into the Political Environment of Government: The Role of Evaluation Policy, *New Directions for Evaluation*, núm. 123, p. 51-66.

- CHELIMSKY, E. (2014). "Public-Interest Values and Program Sustainability: Some Implications for Evaluation Practice". *American Journal of Evaluation*. Vol. 35(4) 527-542.
- COMMISSION OF THE EUROPEAN COMMUNITIES (2007). *Responding to Strategic Needs. Reinforcing the Use of Evaluation*. SEC (2007) 213. Brussels, 21 February 2007. <http://bit.ly/1A5P9ZG>
- CONEVAL (2013). *Diagnóstico para conocer el avance en monitoreo y evaluación en las entidades federativas 2013*. <http://bit.ly/1AYcyNz>
- CONEVAL (2014). *Informe de Evaluación de la Política de Desarrollo Social 2014*. <http://bit.ly/1u0970z>
- COOKSY, L. J.; MARK, M. M.; TROCHIM, M.K. (2009). "Evaluation policy and evaluation practice: where do we go from here", *New Directions for Evaluation*, núm. 123, pp. 103 – 110.
- CRÉMIEUX, L.; SANGRA, E.; (2014). *Evaluation at the Swiss Federal Audit Office*.
- CUNILL-GRAU, N., & OSPINA, S. M. (2012). "Performance measurement and evaluation systems: Institutionalizing accountability for governmental results in Latin America". In S. Kushner & E. Rotondo (Eds.), *Evaluation voices from Latin America*. *New Directions for Evaluation*, 134, 77–91.
- DELEAU, M., (1986). *Évaluer les politiques publiques: méthodes, déontologies, organisation: rapport*. Paris: Documentation française.
- DEPARTAMENTO NACIONAL DE PLANEACIÓN. (2010). *15 años del Sistema Nacional de Evaluación de Gestión y Resultados – SINERGIA*. DNP y Banco Mundial. Bogotá, D.C., Colombia.
- ESV (2013). *Financial information of good quality, in time and for multiple uses – the Swedish experience*. <http://bit.ly/1E0qdEr>
- EUROPEAN COMMISSION (2014). *Programming Period 2014-2020. Monitoring and Evaluation of European Cohesion Policy*. *European Social Fund*. September 2014. <http://bit.ly/1zWqEOZ>
- FEINSTEIN, O.; ZAPICO-GOÑI, E. (2010). "Evaluation of Government Performance and Public Policies in Spain". ECD Working Paper Series, núm. 22, May 2010. IEG. The World Bank Group.
- FORS, K.; REBIEN, C. (2014). "Is there a Nordic/Scandinavian Evaluation Tradition?". *Evaluation*, Vol. 20(4) 467–470.
- FURUBO, J-E. (2013). "When Incremental Change isn't Enough", en Rist, R., Boily, M-H, Martin, F.R (Eds.) (2013); *Development Evaluation in Times of Turbulence. Dealing with Crises that Endanger our Future*. The World Bank
- FURUBO, J.; RIST, R.; SANDAHL, R. (Eds) (2002). *International Atlas of Evaluation*. Transaction Publishers. New Brunswick.
- BEMELMANS-VIDEC, M.L. (2002). "Evaluation in the Netherlands 1990 – 2000: Consolidation and Expansion", en Furubo, Rist y Sandahl (Eds) (2002), págs. 91-114.
- DERLIEN, H-U., (2002); "Policy Evaluation in Germany: Institutional Continuation and Sectoral Activation", en Furubo, Rist y Sandahl, (Eds) (2002), págs. 77-91.
- FONTAINE, c.; MONNIER, E.; (2002) "Evaluation in France", en Furubo, Rist y Sandahl (Eds.) (2002) págs. 63-75.
- FURUBO, J.; SANDAHL, R.; (2002); "A Diffusion Perspective on Global Developments in Evaluation", en Furubo, Rist y Sandahl, (Eds.) (2002) págs. 1-23.
- FURUBO, J.; SANDAHL, R.; (2002); "Coordinated Pluralism – The Swedish Case", en Furubo, Rist y Sandahl, (Eds.) (2002), págs. 115-128.
- GALLEGO, R. (2013). "Decision-making and policy shift under quasi-federal constraints: The case of health management in Catalonia, 2003-2007". Paper presentado en la *63rd Political Studies Association (PSA) Annual International Conference*. 25 - 27 Marzo 2013, Cardiff.
- GALLEGO, R.; SUBIRATS, J. (2012). "Spanish and Regional Welfare Systems: Policy Innovation and Multi-level Governance". *Regional and Federal Studies* Vol. 22, No. 3, 269–288, July 2012.

- GURGAND, M., VALEDENAIRE, M. (2012). «Le Fonds d'Expérimentation pour la Jeunesse», en *Éducation & Formations* no. 81. <http://goo.gl/mwKuM>
- GUZMÁN, M.; IRARRÁZAVAL, I.; DE LOS RÍOS, B. (2014). "Monitoring and Evaluation System: The Case of Chile 1990–2014". ECD Working Paper Series, No. 29, July 2014. IEG. The World Bank Group.
- HERNÁNDEZ LICONA, G. (2009). *Evaluación de la Política Social: Un cambio institucional en México*.
- HERNÁNDEZ LICONA, G.; MARTÍNEZ MENDOZA, E.; GARZA NAVARRETE, T. (2013). *Construyendo un sistema de evaluación del desempeño para el desarrollo social*.
- HOLM, G. (2006). "Evaluation in the Nordic Countries". *Journal of Multi-Disciplinary Evaluation* (JMDE:4)
- INNERARITY, D. (2015). "La decepción democrática", diario *El País*, 2 de febrero.
- JACOB, S. (2005). *Institutionaliser l'Évaluation des Politiques Publiques. Étude comparée des dispositifs en Belgique, en France, en Suisse et aux Pays-Bas*. P.I.E. Peter Lang.
- JACOBI, L.; KLUVE, J. (2006). *Before and After the Hartz Reforms: The Performance of Active Labour Market Policy in Germany*. Discussion Paper Series. IZA. April 2006.
- JACOB, S.; SPEER, S.; FURUBO, J-E. (2015). "The institutionalization of evaluation matters: Updating the International Atlas of Evaluation 10 years later". *Evaluation* 2015, Vol. 21(1) 6–31.
- KICKERT, W. (2011). "Distinctiveness of administrative reform in Greece, Italy, Portugal and Spain. Common characteristics of context, administrations and reforms". *Public Administration*, Vol. 89, No. 3, 2011 (801–818).
- KUHLMANN, S.; BOGUMIL, J.; GROHS, S. (2008). "Evaluating Administrative Modernization in German Local Governments: Success or Failure of the New Steering Government?". *Public Administration Review*; Sep/Oct 2008; 68, 5; pg. 851.
- KUHLMANN, S. (2010). "Performance Measurement in European local governments: a comparative analysis of reform experiences in Great Britain, France, Sweden and Germany". *International Review of Administrative Sciences*, Vol. 76(2):331–345.
- LÁZARO, B. (2012a). "Estrategias para el avance de la evaluación: la experiencia catalana a través de Ivalua". *Presupuesto y Gasto Público*. 68/2012: 25–39. Secretaría de Estado de Presupuestos y Gastos. Instituto de Estudios Fiscales.
- LÁZARO, B. (2012b). "Las nuevas formas de inversión social como motor de innovación de las políticas públicas", en *Avaluació per al Bon Govern*, Noviembre de 2012. <http://bit.ly/160vYTJ>
- LEEuw, F.L. (2009). "Evaluation Policy in the Netherlands". *New Directions in Evaluation*. Special Issue: Evaluation Policy and Evaluation Practice. Issue 123, pages 87–102.
- LEEuw, F.L.; FURUBO, J. (2008). "Evaluation Systems. What are They and Why Study Them?" en *Evaluation*. Vol. 14(2), pp. 157–170.
- MACKAY, K. (2007). *How to Build M&E Systems to Support Better Government*. The World Bank.
- MARK, M.M.; COOKSY, L.J.; TROCHIM, M.K. (2009). "Evaluation Policy: An Introduction and Overview", *New Directions for Evaluation*, núm.. 123, pp. 3–12.
- MAY, E.; SHAND, D.; MACKAY, K.; ROJAS, F.; SAAVEDRA, F., (Eds), (2006). *Hacia la institucionalización de los sistemas de monitoreo y evaluación en América Latina y el Caribe*. Banco Mundial. Banco Interamericano de Desarrollo. Washington D.C.
- NATIONAL AUDIT OFFICE (2013). *Evaluation in Government*. Cross-government Report. December 2013. United Kingdom.
- NEIROTTI, N. (2012). "Evaluation in Latin America: Paradigms and practices". In S. Kushner & E. Rotondo (Eds.), *Evaluation voices from Latin America*. *New Directions for Evaluation*, 134, 7–16.

- NETHERLANDS COURT OF AUDIT (2012). *Evaluation of policy effectiveness in central government*. May 2012. Resumen en inglés disponible en <http://bit.ly/16jxzw>
- NETHERLANDS COURT OF AUDIT (2013). *Evaluation of Policy Effectiveness in Central Government: Follow-up audit*. June 2013. Resumen en inglés disponible en <http://bit.ly/1DniwnE>
- OECD (2009). "Regulatory Impact Analysis. A tool for policy coherence". *OECD Reviews of Regulatory Reform*.
- ONGARO, E. (2009). *Public Management Reform and Modernization. Trajectories of Administrative Change in Italy, France, Greece, Portugal and Spain*. Edward Elgar.
- OSBORNE, D.; GAEBLER, T. (1992). *Reinventing Government*. Addison-Wesley Publ. Co.
- PARRADO, S. (2008). "Failed policies but institutional innovation through 'layering' and 'diffusion' in Spanish central administration". *International Journal of Public Sector Management*, Vol. 21 No. 2, 2008 pp. 230-252.
- POLLITT, C.; BOUCKAERT, G. (2004). *Public Management Reform. A Comparative Analysis*. Oxford University Press (2nd. Ed.).
- RIST, R. (2006): "The 'E' in Monitoring and Evaluation-Using Evaluative Knowledge to Support a Results-Based Management System", From Studies to Streams. Transaction Publishers. New Brunswick.
- RIST, R., BOILY, M-H, MARTIN, F.R (EDS.) (2013). *Development Evaluation in Times of Turbulence. Dealing with Crises that Endanger our Future*. The World Bank.
- RIST, R.; STAME, N. (Eds.) (2011). *From Studies to Streams: Managing Evaluative Systems*. Comparative Policy Evaluation, Vol. XII. Transaction Publishers. New Brunswick.
- RUTTER, J.; HARRIS, J. (2014). *The Special Ones. How to make central government units work*. Institute for Government.
- SIGMA (2014). *The Principles of Public Administration*. <http://bit.ly/1qpny18>
- STAME, N., (2008). "The European Project: Federalism and Evaluation", en *Evaluation*, Vol 14(2): 117 –140.
- STOCKDILL, S., BAIZERMAN, M., COMPTON, D.; eds. (2002). "Toward a Definition of the ECB Process: A Conversation with the ECB Literature", *New Directions for Evaluation*, núm. 93, pp. 7-25.
- TALBOT, C. (2010). "Performance in Government. The Evolving System of Performance and Evaluation Measurement, Monitoring, and Management in the United Kingdom". *ECD Working Paper Series*. No. 24. IEG World Bank.
- TROCHIM, W.M.K. (2009). "Evaluation Policy and Evaluation Practice", *New Directions for Evaluation*. núm. 123, pp. 13-32.
- VAN THIEL, S.; LEEUW F. L. (2002). "The Performance Paradox". *Public Performance Review*, Vol 25 núm.. 3, pp. 267-281
- VARONE, F.; JACOB, S., (2004). «Institutionnalisation de l'évaluation et nouvelle gestion publique: un état des lieux comparatif». *Revue internationale de politique comparée*. 2004/2 - Vol. 11 pages 271 à 292.
- VARONE, F.; JACOB, S.; DE WINTER, L. (2005). "Polity, Politics and Policy Evaluation in Belgium". *Evaluation*, Vol. 11 (3), pp. 253-273.
- VEDUNG, E.; (2010). "Four Waves of Evaluation Diffusion". *Evaluation*. 16(3) 263–277.
- VILLARREAL NAVARRO, J. (2007). "Colombia y el Sistema de Monitoreo y Evaluación. SINERGIA". *Fortalecimiento de los sistemas de monitoreo y evaluación (M&E) en América Latina*. Banco Mundial y Centro Latinoamericano de Administración para el Desarrollo (CLAD).
- VIÑAS, V. (2009). "The European Union's Drive Towards Public Policy Evaluation. The Case of Spain". *Evaluation*, Vol. 15(4), pp. 459-472.
- WOLLMANN, H. (Ed.) (2003). *Evaluation in Public Sector Reform. Concepts and Practice in International Perspective*. Edward Elgar.
- ZALTSMAN, A. (2006). "Experience with Institutionalizing Monitoring and Evaluation Systems In Five Latin American Countries: Argentina, Chile, Colombia, Costa Rica and Uruguay". *ECD Working Paper Series*, No. 16. IEG, Word Bank.

Anexo 2. Referencias consultadas por país y tipo de fuente

Fuente: Publicaciones de órganos del Ejecutivo responsables de la evaluación

- Chile: ARENAS, BERNER (2010)
- Colombia: DEPARTAMENTO NACIONAL DE PLANEACIÓN. (2010)
- México: CONEVAL (2013), (2014)
HERNÁNDEZ LICONA (2009)
HERNÁNDEZ LICONA, MARTÍNEZ MENDOZA, GARZA NAVARRETE, (2013)
- Suecia: ESV (2013)

Fuente: Publicaciones de organismos de control del Ejecutivo

- P. Bajos: NETHERLANDS COURT OF AUDIT (2012), (2013)
- R. Unido: NATIONAL AUDIT OFFICE (2013)

Fuente: Publicaciones de organismos internacionales

- Chile: GUZMÁN, IRARRÁZVAL, DE LOS RÍOS, (2014)
ZALTSMAN (2006)
MACKAY (2007)
MAY, SHAND, MACKAY, ROJAS, SAAVEDRA (2006)
- Colombia: ZALTSMAN (2006)
MACKAY (2007)
MAY, SHAND, MACKAY, ROJAS, SAAVEDRA (2006)
VILLARREAL NAVARRO (2007)
- España: FEINSTEIN, ZAPICO-GOÑI (2010)
- México: MAY, SHAND, MACKAY, ROJAS, SAAVEDRA (2006)
- R. Unido: TALBOT (2010)

Fuente: Artículos de expertos participantes en los sistemas en revistas independientes

- España: ARGILÉS 2014
LÁZARO (2012a)
- Italia: BASSANINI (2009)
- P. Bajos: LEEUW (2009)

Fuente: Artículos de expertos en publicaciones independientes

- Alemania: DERLIEN (2002)
SPEER (2012)
JACOBI, KLUVE (2006)
KUHLMANN, BOGUMIL, GROHS (2008)
KUHLMANN (2010)
WOLLMANN (2003)
JACOB, SPEER, FURUBO (2015)
- Bélgica: JACOB (2005)
VARONE, JACOB, DE WINTER. (2005)
- Chile: CUNILL-GRAU, OSPINA (2012)
- Colombia: CUNILL-GRAU, OSPINA (2012)
- España: VIÑAS (2009)
JACOB, SPEER, FURUBO (2015)
- Francia: FONTAINE, MONNIER (2002)
GURGAND, VALEDENAIRE (2012)
JACOB (2005)
KUHLMANN (2010)
JACOB, SPEER, FURUBO (2015)
- Italia: STAME (2012)
WOLLMANN (2003)
JACOB, SPEER, FURUBO (2015)
- México: CUNILL-GRAU, OSPINA (2012)
- P. Bajos: BEMELMANS-VIDEC (2002)
JACOB (2005)
WOLLMANN (2003)
JACOB, SPEER, FURUBO (2015)
- R. Unido: KUHLMANN (2010)
RUTTER, HARRIS, (2014)
WOLLMANN (2003)
JACOB, SPEER, FURUBO (2015)
- Suecia: BOHNI NIELSEN, WINTHER (2014)
FORS, REBIEN (2014)
FURUBO, SANDAHL (2002)
HOLM, (2006)
KUHLMANN (2010)
JACOB, SPEER, FURUBO (2015)

Consortio Liderado por



Socios Coordinadores



Participan más de 80 Socios Operativos y Entidades Colaboradoras de Europa y América Latina

EUROsocial es un programa de cooperación regional de la Unión Europea con América Latina para la promoción de la cohesión social, mediante el apoyo a políticas públicas nacionales, y el fortalecimiento de las instituciones que las llevan a cabo. EUROsocial pretende promover un diálogo euro-latinoamericano de políticas públicas en torno a la cohesión social. Su objetivo es contribuir a procesos de reforma e implementación en diez áreas clave de políticas, en ciertas temáticas, seleccionadas por su potencial impacto sobre la cohesión social. El instrumento del que se dota es el de la cooperación institucional o aprendizaje entre pares: el intercambio de experiencias y la asesoría técnica entre instituciones públicas de Europa y de América Latina.

