

*Problemática de la actividad evaluadora en el sector público (1)*

Rodney DOBELL y David ZUSSMAN: «Un sistema di valutazione per l'apparato dello Stato: se la politica e teatro, la valutazione e (per la maggior parte) un'arte.»

John NALBANDIAN: «Valutazione della prestazione: se solo non vi fossero coinvolte le persone.»

Jerome A. MARK: «Metodi di misurazione della producttività nel settore pubblico (a livello federale, statale, locale).»

H. G. ROGERS, M. A. ULRICH e K. L. TRAVERSY: «La valutazione dei programmi di intervento pubblico: lo stato delle arti nei governi canadesi.»

Tony BOVAIRD: «Sviluppi recenti delle tecniche di misurazione del prodotto nelle amministrazioni locali.»

La revista *Problemi di Amministrazione Pubblica*, en su número 3 de 1982, ofrece la traducción al italiano de una selección de artículos que tratan, bajo diferentes puntos de vista, un tema tan interesante como actual: la evaluación en el sector público.

Esta serie de artículos representan la opinión de varios profesores y altos cargos de la Administración de los Estados Unidos y Canadá sobre la situación actual de las diferentes líneas de investigación en este campo, los sistemas y técnicas de evaluación implantadas en los distintos niveles

del gobierno (federal, estatal y local) de sus países y la posibilidad de su desarrollo futuro.

La lectura de esta publicación sugiere, entre otras reflexiones que veremos al comentar cada uno de los artículos, que el grado de utilización de las técnicas modernas de evaluación está en una fase inicial y transitoria y que los sistemas de información y la voluntad política juegan el papel más importante en el futuro inmediato del desarrollo e implantación de dichas técnicas.

Sin otro comentario, y como corresponde a nuestro objetivo, pasemos a comentar cada uno de los artículos por separado.

---

(1) «Problemi di Amministrazione Pubblica» núm. 3. Napoli, 1982, 169 pp.

**Sistema de evaluación para el aparato estatal: si la política es el teatro, la evaluación (en gran medida) es un arte (2)**

El término o concepto *evaluación* está llamado a jugar un papel muy importante en el ámbito de la ciencia de la Administración pública. Sin embargo, por el momento existe un cierto escepticismo sobre la utilidad de las técnicas de evaluación de alternativas en el sector público.

Rodney Dobell y David Zussman, profesores de Administración pública en la Universidad de Victoria, Canadá, presentan en su artículo las posibles causas de esta *atmósfera de desilusión* con respecto al análisis y evaluación de alternativas en el sector público.

La primera razón expuesta es la ausencia de un marco teórico y de un objetivo que fuera ampliamente aceptado por los estudiosos y usuarios del desarrollo de las técnicas de evaluación. El estado del conocimiento es fragmentario. Por un lado, psicólogos y sociólogos, desde finales de los años sesenta, vienen utilizando métodos experimentales o cuasi-experimentales de contrastación para llevar a cabo la evaluación de los programas públicos que el Gobierno de los Estados Unidos viene desarrollando desde finales de la década de los sesenta, años en que el presidente Johnson iniciaba su plan político de transformar Estados Unidos en la *Gran sociedad*.

Su método se basa en la comparación de dos grupos o comunidades de

características muy similares. Uno de los grupos, el experimental, es afectado por el programa implantado; el otro, el grupo de control, queda fuera del campo de su acción. Cumplidos estos requisitos y observadas ciertas características de los dos grupos en fechas anteriores y posteriores a la intervención pública, se podrá evaluar el efecto debido única y exclusivamente al programa.

El mayor grado de éxito y fiabilidad de la evaluación depende de que la única diferencia existente entre los dos grupos sea el sometimiento o no a los efectos de la intervención pública.

Sin embargo, en opinión de R. Dobell y David Zussman, esa equiparación entre el grupo de experimentación y el grupo de control resulta imposible de conseguir. De hecho, la mayoría de las evaluaciones se basan en cuasi-experimentos cuya validez metodológica, al no contrastar en el tiempo o con el grupo de control, está muy por debajo de la del análisis experimental.

Por otro lado, las ciencias administrativas y económicas están desarrollando técnicas de evaluación que poseen un fundamento teórico muy sólido y una metodología muy depurada. La técnica más conocida dentro de este enfoque es el análisis coste-beneficio que, en líneas muy generales, consiste en la comparación de costes y beneficios sociales actualizados, que *a priori* se estima ocasionarán los distintos programas o estrategias para la consecución de un objetivo dado. El análisis coste-beneficio permite asignar óptimamente los recursos y tomar decisiones racionalmente al elegir la alternativa óptima entre las presenta-

(2) Rodney DOBELL y David ZUSSMAN: «An evaluation system for government: if politic is theatre, then evaluation is (mostly) art», *Canadian Public Administration*, núm. 3, 1981, páginas 404-427.

das como posibles para la consecución de los objetivos marcados.

El problema radica en la dificultad de cuantificar ciertos costes de carácter social y de especificar claramente los objetivos y finalidad de los programas.

Dobell y Zussman apuntan una tercera tendencia o vía de evaluación que puede considerarse de naturaleza análoga al análisis de mercado. Su método es ecléctico y su marco teórico varía según el destinatario del informe del estudio. Cada evaluación se enfoca de una forma particular, por lo que los resultados no tienen validez externa, no son generalizables.

Dobell y Zussman manifiestan la existencia de dificultades de naturaleza pragmática más que teórico-metodológica. El proceso de recogida y selección de información, la resistencia de la propia estructura político-organizativa que lleva a cabo los programas evaluados y otras limitaciones de carácter intrínseco al análisis, dificultan también la implantación de cualquier sistema de evaluación de programas públicos.

Los autores mencionados proponen como posible solución una doble estrategia de evaluación, o si se prefiere, estrategia de evaluación permanente: *evaluación formativa* y *evaluación definitiva*. La primera supone una verificación del programa en curso y una retroalimentación que permita al responsable del mismo reasignar sus recursos y mejorar el grado de consecución final de los objetivos. La segunda supone un análisis general de los resultados obtenidos tras la terminación total del programa. A diferencia de la anterior, el destinatario del informe de evaluación definitiva no

es director del programa evaluado, sino el Consejo de Ministros o el propio Parlamento.

También R. Dobell y D. Zussman consideran la necesidad de un relanzamiento de las técnicas de evaluación y una mejora en su calidad, basando ésta en un previo desarrollo de los sistemas de información y estadísticas.

Los profesores canadienses terminan afirmando que el análisis y la evaluación son técnicas más propias de un *aula de justicia que de un aula universitaria*. No merece la pena, en su opinión, continuar discutiendo la problemática metodológica. Ha llegado el momento de pasar a la práctica.

#### **Evaluación de tareas: si sólo las personas no estuvieran envueltas (3)**

John Nalbandian, profesor adjunto de la Universidad de Kansas, defiende en este artículo la sustitución de criterios subjetivos por técnicas y métodos objetivos en la evaluación de tareas. Sin embargo, esto no debe suponer la implantación de unas reglas rígidas de comportamiento que paralicen la organización, en lugar de aumentar su eficacia.

Las valoraciones subjetivas facilitan el favoritismo, la ineficacia y los conflictos entre el personal directivo y el personal subordinado. Por otro lado, las técnicas modernas de evaluación de tareas, y, entre ellas, los contratos de actividades (*performance contracts*), intentan evitar estos abusos y al mismo tiempo permiten que los empleados conozcan de forma más clara y concisa las expectativas de la

(3) John NALBANDIAN: «Performance appraisal: if only people were not involved», *Public Administration Review*, núm. 3, 1981, páginas 392-396.

organización con respecto a ellos. Dicho en otras palabras, la organización puede describir de forma explícita, a través de los contratos u otras técnicas, la contribución que ha de aportar el empleado a cambio de la remuneración y otras contraprestaciones que aquélla le ofrece.

Además de estas ventajas, estas técnicas modernas de evaluación de tareas proporcionan elementos o herramientas de gestión muy útiles a la Dirección de las organizaciones: énfasis en objetivos a cumplir, estándares de trabajo y comportamiento relacionado a los resultados. Todo ello debería facilitar un aumento de la productividad en la organización que se implante.

Sin embargo, hay importantes costes inmersos en esta estrategia de evaluación de tareas basada en la objetividad que deben ser tenidos muy en cuenta si se quiere un resultado positivo tras su implantación. En efecto, ciertos comportamientos y actividades realizadas por el personal con carácter informal y de naturaleza no medible pueden ser muy beneficiosos a la hora de obtener una mejora de la calidad del trabajo y la productividad de la empresa. Estas actividades afectan generalmente al sistema de comunicación y coordinación del comportamiento interno y externo de cualquier grupo dentro de una organización.

La inhabilidad de los contratos de actividades para captar todos los comportamientos que puedan mejorar la eficacia de la organización ofrece a los empleados una puerta abierta para actuar únicamente de acuerdo con los criterios por los cuales pueden ser evaluados.

Asimismo, aunque estas técnicas suponen, por un lado, una ayuda a la alta gestión, también, por otro lado, puede plantear límites a la discrecionalidad del *manager*. En concreto, éste puede verse predestinado a una política rígida de retribuciones, promoción y movimiento de personal en general. Según estos argumentos es aceptable pensar que esta estrategia podría incluso reducir la eficacia de la organización.

Por lo tanto, concluye el profesor Nalbandian, no se debe dejar de prestar atención a los factores humanos dentro del proceso de evaluación. Hay varios elementos que son fundamentales para una fructífera implantación de un sistema objetivo de evaluación. En primer lugar, motivación por parte de los supervisores para evaluar, esto es, aceptación del papel evaluador. Segundo, confianza de los subordinados con respecto a las actividades de evaluación. Finalmente, preparación del personal supervisor y establecimiento de una carrera profesional que tenga presente la habilidad, imaginación, tolerancia y conocimiento del mundo interno de la evaluación de actividades por parte del personal a la hora de ser promocionado a puestos de dirección y mando.

#### **Métodos de medida de la productividad en el sector público (a nivel federal, estatal y local) (4)**

En este artículo, Jerome A. Mark expone diversos métodos de cálculo de productividad utilizados por el Ministerio de Trabajo y otros organis-

(4) Jerome A. MARK: «Measuring Productivity in Government: Federal, State and Local», *Public Productivity Review*, marzo de 1981, pp. 21-44.

mos públicos del Gobierno de los Estados Unidos.

Los intentos de evaluar la productividad se pueden clasificar en tres amplios grupos, según estén basados en medidas de eficiencia, medidas de eficacia o medidas operativas. Cada uno de los tres criterios juegan un papel importante en la evaluación de la productividad. Sin embargo, J. Mark, delegado adjunto para Productividad y Tecnología en el Bureau of Labor Statistics, argumenta que gran parte de la investigación sobre los criterios de productividad se centran en el análisis de la medida de eficiencia.

El Bureau of Statistics ha desarrollado un índice de productividad para cada uno de los sectores del Gobierno. Este trabajo lo ha realizado en colaboración con la General Accounting Office, la Office of Management and Budget y la Civil Service Commission, entre otros organismos públicos.

Durante los años setenta se recogieron datos sobre 350 entes gubernativos, aproximadamente el 65 por 100 del empleo público federal. Del resultado del proceso de esta información se ofrecen dos tablas numéricas que muestran la evolución anual, desde 1971 hasta 1978, de los índices de producción por unidad de factor de trabajo empleado y la reagrupación de éstos por categorías funcionales con características comunes. De esta última reagrupación funcional se ofrece también un análisis para los años 1967 a 1978 de la tasa media de incremento anual de la producción por empleado-año dentro del ámbito de la población civil ocupada en la Administración civil. A juicio de J. A. Mark, los cálculos presentados en es-

tas tablas no representan exactamente la productividad del Gobierno federal, sino más bien la media de la variación de productividad de las organizaciones federales.

Según el propio autor, en la medición de la productividad existen algunos problemas de cierta importancia. Dado que la mayor parte de la actividad pública tiene la naturaleza de la producción de servicios, resulta muy difícil definir y cuantificar el producto. Asimismo, si los indicadores de los productos miden una cantidad no homogénea y la unidad administrativa, a distinto nivel de trabajo, cambia en el tiempo la medida de la producción, éstos reflejarán no sólo un cambio en la productividad, sino también en los distintos tipos de producto.

Entre otros problemas, J. A. Mark cita también que, en algunos casos, los indicadores en cuestión no reflejan los cambios en la calidad de la producción.

El principal remedio, según el alto comisionado del Departamento de Trabajo, consiste en implantar un sistema de recogida de datos e información mucho más detallado que permita la mejora de la precisión de las medidas.

Posteriormente, J. A. Mark analiza otros programas o *tentativas* de evaluación de la productividad llevadas a cabo por el Naval System Command, el Veteran's Administration (organismo encargado de gestionar todos los programas relacionados con veteranos de guerra), el Veterinary Service del Departamento de Agricultura, etc.

Como dirección futura, el autor del artículo proclama la necesidad de que el Bureau of Labor Statistics continúe reforzando y mejorando las medidas de producción existentes. Para ello

sería preciso que en una primera etapa se modificara el sistema de registro de datos con el fin de disponer de indicadores más válidos y eficaces, y, así, poder iniciar la medición de producciones que en la actualidad no están sometidas a evaluación. J. A. Mark sugiere la necesidad de elaborar un indicador de carácter general que mida globalmente la producción total de un solo ente. Asimismo califica de imprecisa la medida de productividad que utiliza en la actualidad el BLS, ya que está basada en un solo factor: el trabajo. En este sentido, J. A. Mark propone que se apliquen tantas medidas de productividad como factores de producción se utilicen en la producción.

Respecto a los métodos de medida de la productividad en los Gobiernos estatales y locales, J. A. Mark llega, tras reflexiones sobre diversas investigaciones e informes, a las siguientes conclusiones. En primer lugar, cuanto más extensa es la jurisdicción territorial más probable que calcule y utilice medidas de eficiencia. Segundo, el uso de estas medidas es irregular y varía de servicio a servicio. Esto quizá sea debido a la diversidad y multiplicidad de los sistemas de información a estos niveles. Por último, la probabilidad de que un ente territorial haga uso de medidas de eficiencia aumenta si disfruta un subsidio o sufre algún tipo de control desde la Administración federal.

El artículo termina proponiendo una ampliación en el campo teórico de los métodos de medición de la productividad. La futura investigación sobre métodos de medición y evaluación deberá profundizar, según el autor, en la búsqueda de medidas que

tengan en cuenta el carácter cambiante de la composición del factor trabajo y que comprenda también dentro de sus límites al factor capital.

#### **Evaluación de los programas de intervención pública: el estado del arte en el Gobierno canadiense (5)**

Este artículo tiene por objeto describir la situación actual de los sistemas de evaluación utilizados por los diferentes niveles de gobierno del Canadá. H. G. Rogers, M. A. Ulrich y K. L. Traversy exponen en él los procedimientos formales de evaluación de los programas públicos, la medida de los resultados de los programas en curso de realización y el control interno sobre los procesos de racionalización.

Como principio básico se considera la evaluación parte integral de la gestión de los programas. Por lo tanto, se hace hincapié en la evaluación de la eficiencia y eficacia de los programas públicos.

Los autores describen de forma correlacionada distintos niveles de gestión pública y la amplia gama de procedimientos de evaluación implantados en Canadá. Así, el Gobierno federal, para un primer nivel de planificación estratégica, esto es, de definición de programas alternativos y asignación de recursos se presentan técnicas y procedimientos como el PEMS o sistema de control y dirección de la política y del gasto. El PEMS supone un análisis socioeconómico de carácter global de la eficacia

(5) H. G. ROGERS, M. A. ULRICH y K. L. TRAVERSY: «Evaluation in practice: the state of the Art in Canadian Governments», *Canadian Public Administration*, núm. 3, 1981, pp. 371-386.

y eficiencia de las actividades que se inician o que actualmente están en curso de realización.

Otro procedimiento o método de análisis utilizado en esta fase es la evaluación objetiva y específica de un programa. Se trata, en este caso, de determinar la adecuación de sus objetivos, aclarar las circunstancias que rodean al programa, su definición concreta y la operacionabilidad de sus resultados. La evaluación de un programa debe responder, principalmente, a las siguientes preguntas:

- 1.<sup>a</sup> ¿Es útil el programa?
  - ¿Son relevantes sus objetivos?
  - ¿Es factible y legal la actividad y los objetivos del programa?
- 2.<sup>a</sup> ¿Cuáles son sus impactos y efectos?
- 3.<sup>a</sup> ¿Ha alcanzado el programa los objetivos previstos?
- 4.<sup>a</sup> ¿Hay alguna forma de mejorar los resultados?

Por último, dentro de esta fase de asignación de recursos, la Administración canadiense utiliza la llamada *revisión política* de los programas o serie de prácticas consistente en el estudio de programas públicos bajo la dirección del Consejo de Ministros.

En un segundo nivel, de gestión de los recursos y control, el Gobierno federal utiliza también el sistema PEMS, pero en un plano operativo, seleccionando actividades y objetivos particulares y estimando el coste de los recursos empleados. Otras técnicas y procedimientos utilizados en este nivel son la medida de resultados, el control interno, el análisis de gestión y el examen o revisión en *base A*.

Todos ellos son términos y técnicas de general conocimiento. Únicamente es necesario aclarar aquí que el llamado examen en base A se basa en un método de naturaleza muy parecida al utilizado por el presupuesto en base cero.

En este segundo nivel, el órgano responsable de la evaluación suele ser el Ministerio de Hacienda, aunque para algunos asuntos éste puede ser sustituido por otro órgano central de gobierno. Finalmente, un tercer nivel operativo y de gestión de los programas requiere análogas técnicas de evaluación. En este caso la evaluación es llevada a cabo por los propios Departamentos.

Finalmente, en el plano provincial H. G. Robers, M. A. Ulrich y K. L. Traversy describen el interés de los gobiernos territoriales por la información sobre los resultados y el desarrollo de los procedimientos sistemáticos de programación y control del gasto. Todas las provincias han implantado en una forma u otra el PPBS o el ZBB.

La información recogida es orientada prioritariamente hacia los recursos o *inputs* utilizados en la ejecución de los programas públicos. En un segundo término se presta atención a sus resultados y, por último, siguiendo el mismo orden de importancia, se recoge información sobre el impacto y los efectos de los programas.

Por lo general, el interés de los gobiernos provinciales radica en la eficiencia y economía de los programas. La mayoría de las provincias han desarrollado un sistema de información financiero, y muchas están desarrollando un sistema de recogida de información sobre los resultados de los programas.

**Reciente desarrollo de las técnicas de medición del producto en la Administración local (6)**

En opinión de Tony Bovaird gran parte de los servicios suministrados por los entes locales son medibles. No obstante, reconoce la existencia de ciertos problemas conceptuales en dicha medición.

T. Bovaird cita, por orden de importancia, los principales obstáculos que se presentan para la utilización y desarrollo de las técnicas modernas de medición de la producción.

En primer lugar, dice Bovaird, existe una preferencia clara de los profesionales por tomar decisiones únicamente en base a juicios basados en conocimientos teóricos y experiencias previas en lugar de apoyarse en medidas objetivas de naturaleza externa.

Por otro lado, existe una grave ausencia de información oportuna y apropiada para los niveles de decisión en la Administración local. Desde diversos puntos de vista, existe gran acopio de datos en esta administración; sin embargo, ésta no es significativa a los efectos de gestión y evaluación.

Por último, y como obstáculo principal, cita el autor la resistencia o incapacidad de la clase política y dirigente de los servicios de manifestar claramente sus objetivos y metas a cumplir. En realidad esta oposición no es tanto a la medida en sí del producto, sino a la claridad de sus funciones y el sometimiento al control ciudadano.

Con el fin de desarrollar un marco teórico de medición, T. Bovaird define producto final de un servicio público como la variación experimentada en el bienestar de los beneficiarios o usuarios de la jurisdicción donde se presta el servicio.

Resulta muy difícil descubrir una medida precisa y apropiada del producto final. Sin embargo, existe toda una serie de medidas que podrían ofrecer una buena estimación de los efectos que el servicio público produce en el bienestar de los usuarios.

Por lo general, estas medidas hacen referencia al nivel de actividad alcanzado por el servicio prestado: ratio profesor/alumno, número de casos atendidos por un asistente social, número de suicidios entre los minusválidos, alumno-hora por materia, etc.

Sin embargo, el problema no está sólo en medir el nivel de actividad alcanzado por el servicio, sino también en determinar la existencia de causalidad. No importa que el nivel de actividad sea muy elevado si no existe una clara relación causa-efecto entre el modelo de producción del servicio, o lo que es lo mismo, la actividad y el producto final del mismo. Es muy importante, por lo tanto, comprobar que, de hecho, el servicio público es el causante del cambio de bienestar experimentado en la jurisdicción donde se ha realizado.

Por último, para analizar el cambio inducido por un servicio en el bienestar colectivo, el autor del artículo introduce en el modelo otros dos conceptos: *el campo de acción*, o número de usuarios afectados por el servicio, y *el impacto* o medida de cambio de bienestar de los mismo.

En el resto del artículo, T. Bovaird dedica su atención al análisis del im-

(6) Tony BOVAIRD: «Recent developments in oupput measurement» en *Local Governments Studies*, núm. 5, 1981, pp. 35-53.

pacto del servicio sobre los usuarios del mismo, a través de la fijación previa de los objetivos en términos operativos, ya sea mediante una cuantificación física, mediante una valoración económica que el servicio implícitamente ofrece al usuario, o mediante una estimación subjetiva.

Finalmente, el autor estudia tres casos particulares de servicios públicos locales: transporte, vivienda y defensa del consumidor, permitiendo de esta forma la identificación en la práctica de las tres estrategias antes mencionadas.

Eduardo ZAPICO GOÑI

