

Enfoques y lineamientos para el seguimiento, la medición y la evaluación del desempeño en programas del sector público*

Gerald E. Caiden y Naomi J. Caiden

La preocupación por lograr altos desempeños en los programas del sector público no es nueva. A lo largo de casi un siglo, políticos, administradores y público en general han mostrado interés en la economía, la eficiencia y la productividad de las organizaciones gubernamentales. Los esfuerzos por medir los resultados de las agencias públicas y por evaluar sus impactos o sus productos pretendían alcanzar metas de largo plazo establecidas por los reformadores en procura de un mejor desempeño gubernamental. En años recientes, al ímpetu en favor de la medición se añadió la urgencia, en la medida en que la confianza en el gobierno estaba declinando, las presiones se intensificaban, se ponía un énfasis creciente en el gerencialismo, y se revisaba el rol de la intervención gubernamental. La medición y la evaluación del desempeño son actualmente parte integral de una agenda de reforma, cuyos elementos primordiales incluyen una mayor flexibilidad para llevar adelante las actividades del gobierno, debilitamiento de las formas tradicionales de dar cuenta de las labores realizadas (*accountability*); devolución, descentralización y privatización; mayor atención a las demandas, conveniencias y escogencias del público, y reconsideración de lo que los gobiernos deberían realmente suministrar, directa o indirectamente.

Este documento responde a la Resolución 50/225 de la Asamblea General de las Naciones Unidas, que formuló una convocatoria con el propósito de sumar esfuerzos dirigidos a fortalecer la administración pública para el desarrollo, formar capacidades, y rediseñar el Estado, para poder enfrentar los cambios globales y tecnológicos. La medición y la evaluación de las actividades del sector público pueden ser vistas como aspectos críticos en relación con los esfuerzos para dotar de mayor capacidad a los

RSP

Revista do
Serviço
Público

Ano 52
Número 1
Jan-Mar 2001

Gerald E. Caiden es egresado de la Escuela de Economía y Ciencias Políticas de Londres, consultor, investigador y administrador de organizaciones internacionales como Banco Mundial y ONU.

Contato:
caiden@usc.edu

Naomi J. Caiden cursó estudios en la Escuela de Economía y Ciencias Políticas de Londres, actualmente es profesora de administración pública en la Universidad del Estado de California en Los Angeles.

Contato:
ncaiden@calstatela.edu

gobiernos; ganar mayor eficiencia, productividad y eficacia; mejorar la transparencia y la responsabilidad (*accountability*); recuperar la confianza del público en las instituciones gubernamentales y contribuir a una reorientación del rol y de las funciones del gobierno. El documento explica los usos y las limitaciones de la medición y evaluación del desempeño en el sector público, así como sus implicaciones prácticas en lo concerniente a la capacidad institucional. Asimismo, sugiere formas con las cuales los gobiernos y las instituciones internacionales podrían responder a los pedidos de asistencia en esta área.

Significación

Durante los años 80 y 90, numerosos países occidentales industrializados pusieron en marcha reformas gubernamentales de largo alcance. Sus razones fueron similares, la disminución del crecimiento económico había provocado alarma, entendiéndose que la continua expansión del sector público era inconveniente para la inversión y la productividad en el sector privado. En tales circunstancias, se pusieron en tela de juicio las premisas comunes concernientes a la responsabilidad de los gobiernos por el bienestar individual. Por su parte, los gobiernos encontraron que con los procedimientos tradicionales era mucho más difícil hacer frente a problemas sociales y económicos de una gravedad sin precedentes — inmigración, envejecimiento, aceleración de la desigualdad, aumento de la delincuencia y del desorden, impactos de la tecnología, desempleo. Ante la resistencia motivada por los aumentos impositivos y la disminución de los ingresos, a lo que se sumó la incidencia de gastos inflexibles, los *déficits* presupuestarios parecieron intratables e hicieron imposible responder a las demandas dirigidas al gobierno, mediante el incremento del gasto. La insatisfacción del público con el gobierno fue en aumento, variando desde una protesta generalizada que tenía como blanco a los funcionarios arrogantes y despilfarradores, hasta los ataques en gran escala respecto a la legitimidad de las funciones gubernamentales.

Tanto por razones pragmáticas como ideológicas, las reformas aparecieron como necesarias, y su diseño contempló principalmente dos propósitos. En primer lugar, se adoptó una orientación general en el sentido de recortar los gastos innecesarios del gobierno, mejorar la eficiencia y la productividad, incrementar la flexibilidad gerencial y la capacidad de respuesta, impulsar una mayor transparencia y responsabilidad, y poner al consumidor en un primer plano, considerando a los ciudadanos como usuarios de los bienes y servicios públicos. En segundo lugar, estas reformas a su vez habían de contribuir con una reevaluación de lo que hacían las organizaciones gubernamentales, y decidir si ellas debían continuar o no

con sus actividades. La cuestión no sólo se planteó en términos de si las organizaciones gubernamentales estaban desempeñando bien las funciones correspondientes, sino si en general ellas deberían estar a cargo de las mismas.

Esta nueva generación de reformas administrativas, por consiguiente, fue mucho más allá del dominio puramente técnico. Las recientes reformas plantearon una serie de complejos dilemas que sólo podían satisfacerse prácticamente (para que las reformas pudiesen tener alguna significación) sobre la base de un tratamiento caso por caso. La ideología, las teorías generales y los principios no hubieran sido suficientes para superar la inevitable resistencia política y burocrática, ni para suministrar justificaciones realistas, plausibles y bien fundamentadas respecto al cambio de reglas, al desmantelamiento o a la reducción de las estructuras organizacionales, a la reevaluación de actividades existentes desde mucho tiempo atrás, y a una transformación firme, sostenida y de largo alcance del rol y del *modus vivendi* de las burocracias gubernamentales. La medición del desempeño, el seguimiento y la evaluación no fueron sólo ornamentales ni tuvieron una significación incidental en las reformas. Fueron (o deberían ser) las palancas de cambio esenciales, hasta tanto se tuviera la evidencia de que la reforma estaba actuando y la seguridad permanente de que los gobiernos estaban operando de manera competente, con eficiencia y efectividad, en función del interés público.

La vinculación de las reformas con la evaluación y la medición del desempeño no fue casual. En el transcurso de su desarrollo, las reformas tomaron una orientación predecible y francamente uniforme. En el marco de un compromiso ideológico global en relación con la desregulación y la reducción del tamaño del gobierno, las reformas enfatizaron la devolución (el desplazamiento de funciones y servicios hacia niveles más bajos del gobierno), la privatización, la descentralización de la gerencia, la transparencia y la satisfacción del usuario (ver *Organisation for Economic Cooperation and Development*, 1996: 15). Todos estos aspectos dependían, para lograr credibilidad y el mantenimiento de la *accountability*, de la medición de los resultados y de la evaluación de los programas. La desregulación exigía una evaluación de la efectividad, de los costos y de las consecuencias de las regulaciones existentes, y el seguimiento de los efectos de su eliminación o su reemplazo por otros medios alternativos de suministro. La devolución, en particular en los casos en que se mantenía al menos en parte algún financiamiento central, requería el seguimiento de los fondos y el monitoreo de resultados. La privatización, al igual que la contratación externa de servicios, reclamaba la necesidad de establecer estándares, de acordar un seguimiento y de medir el desempeño. Cuando se aplicaban tarifas al usuario, era necesario examinar el costo real de los

servicios y su valor en el mercado. La descentralización de la gerencia, que involucraba el relajamiento de los controles tradicionales sobre el presupuesto y el personal, en beneficio de la flexibilidad gerencial y de la orientación empresarial, implicaba una responsabilidad gerencial concomitante con respecto a resultados que tenían que ser medidos. En los casos en que el gobierno procedió a crear agencias cuasi-independientes, manejadas esencialmente sobre la base de principios comerciales, las medidas de desempeño eran esenciales para asegurar que estas agencias verdaderamente actuaban de conformidad con los objetivos cuantitativos que habían sido establecidos para ellas. La transparencia y la satisfacción del consumidor requerían de la publicación de estándares objetivos mediante los cuales el público pudiese juzgar si las agencias del gobierno se estaban desempeñando satisfactoriamente, y aún más, si estaban haciendo un buen trabajo.

Todos estos aspectos de las reformas recientes cobran mayor relieve en tanto conceptos de la capacidad de gestión de gobierno (*governance*), destinados a sustituir las concepciones más tradicionales acerca de cómo deberían operar los gobiernos (*Organisation for Economic Cooperation and Development*, 1996: 8). Un componente fundamental de la noción de capacidad de gestión de gobierno (*governance*) está en la idea de asociación y cooperación entre los sectores público, privado y sin fines de lucro. En esta asociación, el gobierno no puede ser más el único y ni siquiera el proveedor dominante de numerosos servicios públicos. Su rol es más bien el de financiador, prestador, contratante, comprador o regulador de servicios suministrados por otros. Por consiguiente, el gobierno se ubica en una posición altamente dependiente de la información proveniente de una variedad de fuentes, que concierne al cumplimiento de los estándares, la sujeción a los requerimientos contractuales, el seguimiento de los fondos, la efectividad en el logro de los objetivos, y la calidad y cantidad de servicios. Pese a que las organizaciones gubernamentales pueden operar con mayor autonomía, los gobiernos siguen siendo responsables en último término de asegurar que las metas públicas se cumplan efectivamente (independientemente de los medios utilizados para lograrlo). A tales efectos, los gobiernos necesitan poder medir y evaluar el desempeño, para utilizar esos datos en la elaboración de las políticas, el financiamiento, la conformación del personal, la asunción de responsabilidad y el control.

El estado del arte

Tanto las medidas del desempeño como la evaluación deberían diseñarse para responder a la interrogante básica de “¿Cómo se están haciendo las cosas?” La evaluación es una investigación exploratoria de

un programa o agencia, llevada a cabo en intervalos periódicos o episódicos. Las medidas del desempeño son estimaciones cuantitativas o cualitativas en el tiempo, acerca de lo que está haciendo una organización, cuán bien se está desempeñando y cuáles son los efectos de sus actividades. Entre las medidas del desempeño más comúnmente utilizadas se incluyen la siguientes:

- . **Insumos** (*inputs*), tales como dinero, personal, equipos, materiales.
- . **Carga de trabajo** (*workload*) o niveles de actividad, tales como solicitudes procesadas, niveles de inventario, inspecciones realizadas, estudiantes en clase.
- . **Resultados** (*outputs*), tales como niños vacunados, millas de carretera construidas, toneladas de desechos recogidos, estudiantes graduados.
- . **Impacto** (*outcomes*), de los productos o servicios, tales como enfermedades prevenidas, porcentaje de impuestos recolectados, niveles de pureza del aire alcanzados, niveles de seguridad laboral alcanzados, alivio de la pobreza.
- . **Productividad** (*productivity*), tales como casos atendidos por especialistas, solicitudes procesadas por persona, llamadas de emergencia canalizadas.
- . **Costos** (*costs*), tales como costos promedio para construir una milla de autopista, educar un niño, mantener una piscina pública u otras instalaciones públicas.
- . **Satisfacción del usuario** (*customer satisfaction*), tales como el número de quejas recibidas, resultados de encuestas, utilización de procesos participativos.
- . **Calidad y oportunidad del servicio** (*service quality and timeliness*), tales como tiempos de respuesta de la policía, capacidad para acceder a una agencia, tiempo de espera, aceptación de los horarios de transporte, tasas de deterioro, disponibilidad del servicio (*United States General Accounting Office*, mayo de 1992: 2).

Allí donde se han especificado los niveles para una determinada medida ellos constituyen los **estándares de desempeño**. Cuando la referencia se hace para identificar e implementar mejores prácticas a efectos de exponer un estándar para la comparación de resultados y el mejoramiento del desempeño, esto se conoce como **benchmarking**. Cuando el resultado no es directamente medible, pueden usarse medidas aproximadas conocidas como **indicadores** (*United Kingdom Audit Office*, junio de 1992: 16). Los **incentivos para el desempeño** pueden utilizarse para estimar el desempeño individual con el propósito de determinar la remuneración apropiada y premiar la productividad.

Estas medidas están diseñadas para apreciar la economía, la eficiencia y la efectividad con la que una organización pública suministra

sus servicios. La **economía** tiene que ver con la capacidad de una organización para operar al más bajo nivel de costos posible, y puede estimarse a través de medidas de los insumos y de comparaciones mediante el *benchmarking*. La **eficiencia** es la relación entre los insumos y los resultados, por ejemplo la utilización de insumos mínimos para producir un resultado dado, o la obtención de los máximos resultados para un determinado nivel de insumos. Puede medirse a través de los resultados, la productividad y la medición de los costos. La **efectividad** marca el éxito del programa y se centra en sus resultados, tal como son evaluados a través de la medición del impacto. Además, la calidad de los servicios puede estimarse a través de medidas del proceso, relacionadas con la satisfacción y las opiniones del usuario.

Hacia comienzos de los años 90, el uso de mediciones del desempeño, en una u otra forma, se había divulgado en el sector público. Las encuestas informaban que más de dos tercios de las ciudades en los Estados Unidos utilizaban sistemas de seguimiento del desempeño, y aproximadamente la mitad de las agencias del gobierno federal declararon que utilizaban esta información como apoyo para la toma de decisiones (Swiss, 1991: 155). En varios países se inició la incorporación de sistemas globales para el mejoramiento de la gerencia. En Australia, el *Program Management and Budgeting* y los *Financial Management Improvement Programmes*, utilizaron la planificación estratégica para cotejar las metas con los resultados y los impactos. En el Canadá, varias reformas concibieron acuerdos sobre el desempeño, con una mayor flexibilidad gerencial para enfatizar el mejoramiento de la calidad de los servicios gubernamentales y alcanzar las metas de desempeño establecidas. Las reformas de Nueva Zelanda enfatizaron los métodos empresariales a través del *State Sector Act*, de 1988, y la *Public Finance Act* de 1989, centrándose en la medición de los resultados. En el Reino Unido, a través de la ambiciosa descentralización del *Next Steps Program* a fines de 1980, se enfatizaron ampliamente los logros del desempeño en términos de resultados. En los Estados Unidos, la *Government Performance and Results Act*, de 1993, ordenó a los departamentos federales que implementaran la planificación estratégica a efectos de establecer las metas y medir los resultados de sus actividades (*United States General Accounting Office*, mayo de 1995: 14-5).

Todas estas experiencias indican que no existe una única forma de realizar las evaluaciones o de escoger las medidas del desempeño (Rosen, 1993: 71). Se ha sugerido que el uso de múltiples medidas podría ser de mayor utilidad para “captar” la efectividad de una organización o de un programa, y que el desarrollo de medidas confiables y válidas no debería depender de los datos inmediatamente disponibles, sino más bien basarse en la escogencia deliberada y en la discrecionalidad. Sería posible elaborar

una “cadena” o “jerarquía” de medidas, comenzando con los insumos y terminando con los productos o los resultados finales. Mientras que al comienzo de la cadena los insumos podrían ser fácilmente obtenidos e identificados, ellos pudieran tener poca relación con la efectividad real de la agencia. Por otra parte, los esfuerzos por medir los resultados podrían dificultarse en razón de su magnitud y de su naturaleza difusa, aunque ellos representen el impacto real derivado de las actividades de la agencia (Swiss, 1991: 147). En suma, es de la gerencia pública que depende la escogencia y el balance de las medidas más útiles y más factibles.

Dicho de otro modo, las medidas y la evaluación del desempeño tienen una variedad de usos. Pueden emplearse para dar cuenta de actividades pasadas, de operaciones en curso, y para evaluar los progresos hacia objetivos planificados (*United States General Accounting Office*, mayo de 1992: 2).

Uno de los usos sistemáticos más antiguos de la evaluación fue con propósitos de contabilidad, específicamente en relación con el proceso presupuestario. Las medidas de desempeño tuvieron que ser incorporadas en las presentaciones presupuestarias de las agencias, en un intento de reemplazar el tradicional proceso presupuestario basado en ítems por el presupuesto basado en el desempeño. Las decisiones acerca de las futuras asignaciones de recursos deberían apoyarse en los datos del desempeño, que revelarían el grado en el cual las agencias habían alcanzado sus metas. De hecho, esta reforma, al igual que otros intentos similares efectuados posteriormente (presupuesto por programa y desempeño, presupuesto de base cero), encontraron serias dificultades, dado que, además de la naturaleza política de las asignaciones presupuestarias, no siempre resultaba claro cómo las medidas del desempeño podrían ser utilizadas para determinar los presupuestos. Esta experiencia parece haber contribuido con la disminución del énfasis en los aspectos presupuestarios de la medición del desempeño que se aprecia últimamente, aunque uno de los principales propósitos es el de mejorar la responsabilidad por los resultados de los programas, incluyendo el fortalecimiento del proceso de auditoría. Sin embargo, el gobierno de los Estados Unidos y la Ley sobre Resultados del Desempeño, de 1993, todavía expresan como un objetivo prioritario el establecimiento de vínculos más estrechos y claros entre los recursos y los resultados (*United States General Accounting Office*, marzo de 1997: 1).

Cada vez más jurisdicciones informan que encuentran de utilidad las mediciones del desempeño para la gerencia de las operaciones habituales y el mejoramiento de la productividad. Ellas pueden promover una toma de conciencia sobre los costos, conformar las bases para la introducción de sistemas de pago vinculados al desempeño, y establecer estándares sobre normas aceptables de trabajo. Igualmente, ellas pueden ser utilizadas

para mejorar la calidad de los servicios públicos, mediante la identificación y la atención de las cambiantes necesidades del público, comparando los desempeños en el tiempo, entre organizaciones y en relación con estándares, y a través del seguimiento y el control de los resultados de políticas y programas. Las mediciones del desempeño deberían también contribuir a mejorar la gerencia de la agencia, en razón del énfasis que ponen en los resultados, en la calidad de los servicios y en la satisfacción del usuario, así como por la planificación para alcanzar los objetivos del programa. Igualmente, deberían ayudar a optimizar la eficiencia, clarificar las decisiones y a hacer pública la información sobre el desempeño del programa. Deberían mejorar también la elaboración de las políticas a nivel legislativo, las decisiones sobre gastos y la supervisión general del programa.

Otra utilización de las medidas del desempeño tiene que ver con la evaluación del trabajo que está siendo llevado a cabo bajo contrato. Los contratos deberían establecer claros estándares de desempeño, tomar las provisiones para el seguimiento y especificar los términos de aceptación. Allí donde un nivel de gobierno provee de financiamiento para que algún otro se encargue de realizar ciertas funciones, también se requiere el establecimiento de estándares de desempeño, y asegurar la realización del seguimiento y los términos del cumplimiento de los mismos.

Pero en años recientes, los defensores de la evaluación y de las medidas del desempeño han ido más allá de considerarlas como un modesto aporte adicional al mejoramiento de la productividad, para percibir las como un aspecto integral de la planificación estratégica, por cuanto podrían ser utilizadas para cotejar las metas y evaluar los progresos hacia el logro de los objetivos planificados. Por ejemplo, en la *United States Government Performance and Results Act*, de 1993, las medidas del desempeño se encuentran estrechamente vinculadas con la planificación estratégica desde una triple perspectiva. En primer lugar, a las agencias se les solicitó que elaborasen un plan estratégico identificando sus metas, incluyendo el impacto de las actividades programáticas. De este plan deben desprenderse planes de desempeño anuales, con el desarrollo de medidas de desempeño indicadoras de los progresos hacia las metas establecidas. Por último, las agencias deben recolectar y analizar los datos requeridos para informar sus resultados en informes de desempeño anuales. De esta manera, se espera hacer de la medición de resultados una parte integral de la planificación y de la formulación de políticas en todos los niveles del gobierno (ver *United States General Accounting Office*, mayo de 1997).

En la actualidad existe una literatura significativa acerca de cómo elaborar medidas del desempeño y conducir una evaluación de las organizaciones públicas, así como también se cuenta con estudios de caso e informes que detallan la experiencia en la materia. No obstante, está

claro que todavía subsisten numerosas dificultades en el desarrollo y la utilización de la evaluación y de las medidas del desempeño. La implementación de medidas de desempeño y de una información para la evaluación que sea de utilidad, requiere algo más que experticia técnica. Es necesario construir capacidades y concitar cooperación para el desarrollo y la utilización de la información sobre el desempeño, de modo que verdaderamente ella pueda influir en la toma de decisiones. Más allá de lo valioso y meritorio que pueda ser el propósito que guía el desarrollo de las mediciones del desempeño, si por cualquier razón ellas no pueden ser implementadas, no habrán de ser de utilidad. La experiencia del pasado sugiere prudencia a este respecto.

Implementación

Las actividades de los gobiernos contemporáneos exigen una nueva dimensión para asegurar la responsabilidad (*accountability*), que va más allá y que difiere del respeto de la legalidad, el cumplimiento de los procesos establecidos, la jerarquía burocrática y la probidad financiera. Aún cuando los gobiernos se retiren de la actividad empresarial, desregulen e intenten recortar servicios, ellos siguen estando involucrados en cuanto a dirigir grandes compañías (correos, defensa, aeropuertos, fondos de pensiones); brindar servicios económicos a los mercados; mantener responsabilidades en el área de la salud, del bienestar y de la educación de grandes poblaciones; asegurar la vigencia de la ley y el orden; mantener y restaurar ambientes seguros y ecológicamente sustentados e invertir en grandes proyectos de infraestructura. Pero las formas en que estas funciones son llevadas a cabo están cambiando y difieren de las prácticas del pasado.

En primer lugar, las restricciones presupuestarias ante las crecientes demandas significan que es necesario encontrar nuevas formas de darles respuesta. Por ejemplo, a medida que las poblaciones envejecen, los costos de atención de la salud crecen, y las formas existentes de atender a los ancianos y a los individuos más vulnerables pueden ser tanto demasiado costosas como inadecuadas.

En segundo lugar, ha habido un reconocimiento de que las formas anteriores de encarar los problemas, exclusivamente por intermedio de las burocracias gubernamentales, pueden no funcionar bien en una compleja sociedad post-industrial. Áreas tales como la educación, la seguridad urbana, la recolección de residuos sólidos, la atención de la población dependiente y la regulación ambiental, requieren de una activa cooperación de los usuarios.

En tercer lugar, la provisión monopólica de servicios por parte del gobierno ha sido vista a menudo como demasiado inflexible, permitiendo

poco margen para la variación de las condiciones o para la escogencia por parte de los usuarios del servicio.

En cuarto lugar, el desarrollo de un movimiento en favor de mayor autonomía para los niveles subnacionales de gobierno, exigiendo mayores responsabilidades en la obtención y en el gasto de los ingresos, trajo como consecuencia que los administradores locales y regionales tengan que manejar organizaciones complejas y multifuncionales, que deben responder ante los ciudadanos y las fuerzas económicas locales, aún cuando esto pueda complicar las tareas de los funcionarios nacionales a la hora de implementar las políticas correspondientes.

En quinto lugar, la nueva tecnología de la información ha modificado las formas de realizar negocios, expandiendo las fuentes de la información y las comunicaciones hacia los ciudadanos comunes, e inaugurando nuevas posibilidades para llevar a cabo en diferentes modalidades la implementación de los programas gubernamentales.

En sexto lugar, en muchos países el nivel de corrupción ha alcanzado un punto en el cual ya no es más tolerable, por lo que se impone la toma de medidas para incrementar la transparencia en las organizaciones públicas, y asegurar que ellas estén llevando a cabo las tareas para las cuales fueron creadas.

Finalmente, con la creciente globalización, los responsables de la elaboración de las políticas públicas tienen que responder a problemas que van más allá de las fronteras de sus jurisdicciones, lo que implica la necesidad de entender las fuerzas globales, la capacidad para responder a las tendencias y eventos internacionales, y la habilidad para establecer vínculos de cooperación con organizaciones más allá de las fronteras nacionales.

Estos desarrollos han modificado lo que los gobiernos hacen y cómo lo hacen. Los gobiernos se encuentran inmersos permanentemente en una serie de escogencias de fines y de medios. Operan en un entorno de incertidumbres complejas y convergentes, y se espera de ellos que tomen decisiones sofisticadas, en respuesta a las múltiples presiones, que se involucren en negociaciones y que asuman riesgos. Casi todas las acciones que los gobiernos toman, los involucran en relaciones con otros — otras agencias, otros niveles de gobierno, compañías privadas, grupos de presión, organizaciones sin fines de lucro y ciudadanos. Para ser efectivos, los gobiernos tienen que estimular la cooperación y la participación y embarcarse en empresas conjuntas, operar agencias cuasi-públicas, hacer contratos, concertar acuerdos y concretar una variedad de arreglos financieros. En esta situación, la rendición de cuentas ya no puede basarse totalmente en la jerarquía burocrática que sólo asegura que se cumplió con los procedimientos legales. En la medida en que se espera que los funcionarios gubernamentales actúen con mayor autonomía para ejercer

su juicio profesional sobre la base de lo cotidiano, y confiar en otros para que lleven a cabo las políticas públicas, se necesitan otros medios para asegurar que las políticas se elaboran y se implementan en pro del interés público, y justificar los programas públicos por sus resultados. ¿Cuál es el enfoque viable para medir y evaluar los programas y las políticas públicas?

En el pasado se tejieron expectativas exageradas en torno a las medidas del desempeño y a la evaluación de programas. No es realista pensar que ellas pueden resolver los problemas de la eficacia gubernamental y del proceso de elaboración de las políticas. Con demasiada frecuencia la atención se centró en las cuestiones técnicas de selección de las medidas correctas y de la recolección de datos. Pero las medidas del desempeño y la evaluación de programas son algo más que información; son parte integral de los procesos de gerencia y del entorno político, aspectos ambos que son cruciales para el éxito de las mismas.

Gran parte de la literatura sobre productividad y gerencia orientada a los resultados se centra en la selección de las medidas del desempeño. Un estudio reciente de la *United States General Accounting Office* expone seis “lecciones claves” a partir de la experiencia de cuatro países encuestados (Australia, Canadá, Nueva Zelanda y el Reino Unido). Cinco de estas lecciones están centradas en la naturaleza de la información requerida: 1) los tipos de medidas del desempeño desarrolladas deberían emanar de los objetivos de los programas, ya sea que estén orientadas hacia el resultado (*output*) o hacia el impacto (*outcome*); 2) las medidas del desempeño deberían reflejar la capacidad de la gerencia del programa para influir en el logro de las metas referidas a resultados o a impactos; 3) el personal del programa debería jugar un papel en el diseño de las medidas de desempeño para su programa; 4) los sistemas de medición del desempeño deberían recolectar datos e informar en torno a unas pocas medidas claves, y mantener un equilibrio en cuanto a mostrar diferentes dimensiones del desempeño, tales como la cantidad, la calidad, la eficiencia y el costo; 5) los sistemas de medición del desempeño deberían incluir información cualitativa, además de información cuantitativa sobre el desempeño, así como interpretaciones de los resultados del desempeño; y 6) los sistemas de medición del desempeño deberían suministrar información más agregada sobre el logro de las metas de la organización a la alta gerencia y una información más detallada a los gerentes de programas (*United States General Accounting Office*, mayo de 1995: 4-5).

Con toda claridad, la escogencia apropiada de medidas del desempeño es crítica para el éxito, pero en modo alguno está exenta de dificultades. En general, la información acerca de “resultados” es deseable, pero puede ser demasiado difícil de recolectar y de interpretar, en tanto que los datos referidos a los insumos y a la carga de trabajo, son vistos como relativamente fáciles de recolectar, pero insuficientes. En Nueva

Zelanda y en el Reino Unido, los gobiernos han requerido medidas de resultados (en oposición a impactos). Aún cuando los gobiernos de Australia y del Canadá enfatizaron el establecimiento de objetivos para los programas y para la medición de sus resultados, se aprecia que las mediciones en los informes sobre desempeño, en su mayoría, también estaban orientadas hacia los productos (Ibid.: 4). No existe un sistema ideal.

El grado de discrecionalidad en la selección de medidas de desempeño y en el establecimiento de un marco de referencia para la evaluación suscita interrogantes acerca de la gerencia. ¿Quién debería estar involucrado en la evaluación de programas? Los evaluadores externos podrían ser objetivos, pero también presentarían la desventaja de su falta de conocimiento del programa. En todo caso, el empleo de los mismos en forma regular podría ser costoso y plantea la eventualidad de la creación de todo un ejército de evaluadores profesionales, ya sea públicos o privados, quienes no contribuirían directamente con los resultados de los programas. Si se recurre a la auto-evaluación, los sesgos son probables: después de todo, es poco probable que los funcionarios escojan medidas que los muestren a ellos mismos desde un ángulo desfavorable. En todo caso, el problema de quién debe hacer la evaluación puede escindir la organización entre evaluadores y evaluados. Si funcionarios de alto nivel quedan a cargo de las actividades de evaluación, esto reforzaría su autoridad, pero también significaría que los funcionarios de más bajo nivel se mostrarían escasamente identificados con la evaluación, que pudieran verla como una amenaza, o dejar de suministrar la información y la cooperación necesarias. Si los funcionarios de rango menor se encargan de la tarea, ellos se verían incentivados a justificar sus propias actividades. Si cada quien hace la evaluación, ¿en qué medida esto no drena recursos de las actividades programáticas en curso? Más aún, si la medición del desempeño se utiliza para otorgar incentivos monetarios por el desempeño, a menos que los criterios sean absolutamente claros en relación con las actividades de rutina, la vía queda abierta para el favoritismo, y es probable que se afecte la moral de los no recompensados.

Por consiguiente, la producción de una información sobre el desempeño que sea de utilidad en relación con los programas, requiere de un alto grado de cooperación en el seno de las agencias, y de una gerencia y un liderazgo competentes. Pero aún siendo relevantes, apropiados y oportunos los datos no cumplirán con propósito alguno si no son verdaderamente utilizados. Entre los usuarios potenciales de la información sobre el desempeño y la evaluación no sólo se incluyen los gerentes de programas y los administradores de alto nivel de las agencias, sino también legisladores, ministros, la agencia financiadora central, los cuerpos de auditoría, los partidos políticos, la prensa y el público en general. Pero si bien todos estos interesados pudieran suscribir los propósitos generales de

la evaluación y la medición del desempeño de los programas, cada uno de ellos tiene sus propias preocupaciones y su propia agenda al respecto. Por ejemplo, los legisladores pueden enfatizar la importancia de la supervisión general de la agencia, el desempeño en el corto plazo, la información consistente y la rendición de cuentas, vale decir, elementos que sustentan sus agendas y sus motivaciones políticas. Por otra parte, los funcionarios ejecutivos pueden enfatizar las metas de largo plazo, la adaptabilidad a las necesidades cambiantes y la flexibilidad en la ejecución (*United States General Accounting Office*, marzo de 1997: 3). Por esta razón, los defensores de la medición del desempeño y de la evaluación sugieren que los objetivos y las medidas del desempeño debieran ser objeto de negociación entre los interesados.

La implementación exitosa de la medición del desempeño y de la evaluación de programas no puede ser supuesta de antemano. La información es contingente; la gerencia debe ejercerse sobre la marcha; y los problemas políticos están siempre presentes. La medición del desempeño y la evaluación de programas son técnicas, pero más que eso, son tanto insumos como resultados de una gestión mejorada. Son parte de un proceso de cambios, y como toda innovación, necesitan de un elevado nivel de apoyo, con una introducción progresiva, apoyada en experiencias piloto, para estimular un proceso de aprendizaje. Es igualmente importante tener una comprensión clara de sus limitaciones.

Limitaciones

El hecho de que se hayan hecho esfuerzos durante aproximadamente cincuenta años para desarrollar y utilizar medidas de desempeño en los programas de gobierno y que todavía se sigan confrontando grandes dificultades para poder hacerlo, debería llamar a la reflexión sobre el problema. Los cuestionamientos más serios se refieren a que las medidas, aún cuando hayan sido desarrolladas, no han sido utilizadas. Por ejemplo, después de casi tres años transcurridos en fases preparatorias, la *United States General Accounting Office* informó en junio de 1997 que las agencias del gobierno federal "... todavía no han desarrollado la información necesaria para determinar si sus programas están obteniendo los resultados deseados. Ni siquiera en aquellas instancias en que las agencias tienen información orientada hacia los resultados se está utilizando la información en una medida significativa a efectos de tomar decisiones que afectan sus programas" (*United General Accounting Office*, junio de 1997: 50). Igualmente, un informe anterior sobre presupuestos basados en el desempeño a nivel de los estados, concluía que: "Los gerentes de programas, en su mayoría, han señalado que ellos no utilizan las medidas

habitualmente publicadas en sus documentos presupuestarios” (*United States General Accounting Office*, febrero de 1993: 7). Varios obstáculos han surgido para la difusión del uso de las medidas de desempeño.

En primer lugar, el desarrollo de medidas significativas del desempeño de programas no es una tarea fácil, y muchas agencias carecen de la capacidad y de la orientación para hacerlo. En particular durante un tiempo de restricciones presupuestarias, puede faltar personal capacitado, tiempo y recursos para una actividad exigente que no contribuye directamente con la operación cotidiana de la organización. La incorporación de los datos sobre el desempeño en la toma de decisiones también depende no sólo de la disponibilidad oportuna de los mismos, sino también de la cultura de la organización, que orienta sus actividades, por ejemplo, de acuerdo con hábitos establecidos, por reglas impuestas, basadas en una obediencia sin apelaciones a los jefes ejecutivos, o con base en reacciones ante eventos externos a medida que ellos ocurren. La utilización de datos sobre el desempeño para la toma de decisiones en estos casos requeriría de una revolución organizacional.

No todas las actividades pueden ser significativamente cuantificadas. Actividades de rutina, tales como el procesamiento de formas, la recogida de la basura, la inscripción de estudiantes, el servicio de comidas y la realización de inspecciones, son actividades que se prestan con facilidad a la medición cuantitativa. En el extremo opuesto de la escala, la diplomacia, el asesoramiento de políticas, las negociaciones, las actividades culturales, la investigación y numerosos servicios sociales no se amoldan a dicha medición. Entre ambos extremos, se despliega toda una zona en la que las medidas del desempeño deberían ser aplicadas tomando en consideración una serie de dimensiones, porque es preciso captar tanto la calidad del servicio como la cantidad del mismo. En esta área se ubican, por ejemplo, la salud pública (vgr. la calidad del aire, el control de enfermedades contagiosas, el bienestar infantil), la seguridad pública (respeto de las leyes de tránsito, limpieza de las calles, cumplimiento de las leyes contra el uso ilícito de drogas), la educación (curiosidad, socialización, recreación, desarrollo), y el bienestar social (guarderías, violencia doméstica, entrenamiento y desarrollo laboral, asistencia pre y post-natal). Estas áreas plantean desafíos para la evaluación, porque involucran considerables variaciones en el comportamiento y en la valoración, que difícilmente son compatibles con medidas estandarizadas.

Aún cuando se adopten y se utilicen buenas medidas del desempeño, ellas representan sólo una dimensión de las actividades de una organización. Una organización debe alcanzar sus metas, pero puede hacerlo de una forma tal que contradiga los valores democráticos y que subordine al logro de sus objetivos toda otra consideración. Por ejemplo, una agencia impositiva puede intimidar a los contribuyentes y tratarlos de manera injusta,

y así recolectar grandes montos de ingresos. La policía puede proclamar que ha reducido la delincuencia, pero hacerlo al costo de la brutalidad, la discriminación, la creación de un clima de temor o del predominio de un poder militarizado en la sociedad civil. Las medidas de desempeño pueden también contribuir con estos resultados, en caso de que los individuos sean recompensados o penalizados esencialmente en función de los resultados, sin importar cuán aceptables son los medios que han utilizado para alcanzar dichos resultados.

Allí donde las medidas del desempeño están siendo desarrolladas explícitamente para contribuir con el proceso presupuestario, surgen dificultades agregadas. No ha quedado claro cómo los cambios en los niveles de financiamiento afectan realmente los resultados de los programas. La relación entre medidas de desempeño y asignación de recursos no es lineal. Si un programa está trabajando mal, y muestra escasos resultados, ¿ello significa que debe ser interrumpido o que debe ser provisto de mayores recursos para que puede hacer un mejor trabajo? A la inversa, si un programa está haciendo las cosas bien y alcanzando sus objetivos ¿debería ser dotado de mayores recursos para que haga un trabajo todavía mejor, o deberían recortarse estos recursos, con base en la presunción de que sus propósitos ya se han cumplido y que el programa ya no es más necesario? ¿Cuál es el parámetro valorativo para decidir si una determinada suma de dinero es demasiado grande o demasiado pequeña, o si es lo justo para preservar una especie, operar un sistema de centros para la atención de emergencias traumáticas, o asegurar el monitoreo o el control de enfermedades contagiosas? Más aún, un programa carente de recursos por razones políticas o por causa de restricciones presupuestarias puede estar condenado al fracaso, independientemente de la validez de sus propósitos y de la capacidad de los encargados de su implementación. Igualmente, el esfuerzo por generar todo el tiempo datos sobre el desempeño para el presupuesto anual, puede resultar en medidas superficiales, de escasa utilidad para los responsables de la toma de decisiones, pero que al mismo tiempo los recarga en su trabajo por su detalle y su volumen. Diferencias de perspectivas entre los legisladores y los funcionarios ejecutivos, han contribuido también con la insatisfacción en torno al uso de las medidas del desempeño. Los legisladores han ignorado los datos sobre el desempeño, porque estos datos exigen un formato no familiar para ellos, que utiliza información de base programática, diferente de la información habitual de las estructuras de contabilidad. Ellos han sentido que el uso de las medidas estimulaba el conflicto y que principalmente servía a los fines de los funcionarios ejecutivos, incrementando la discrecionalidad de los mismos. Por su parte, los funcionarios ejecutivos han desarrollado mediciones del desempeño con el propósito de satisfacer necesidades

internas de sus propias agencias y como justificaciones para las formulaciones presupuestarias iniciales, sin preocuparse por los usuarios externos (*United States General Accounting Office*, febrero de 1993: 6-7).

El desarrollo de las medidas del desempeño también se ha visto dificultado en lo que se refiere a especificar las medidas del impacto. Para numerosas actividades, la medición del impacto no parece ser apropiada y no todo, puede ser cuantificado en forma significativa, por ejemplo, las actividades de prevención y de erradicación o aquellas actividades que se refieren a eventos extraordinarios. A menudo las agencias son capaces de sintetizar las actividades de sus programas, pero han tenido problemas al pasar de los resultados a los impactos y han surgido controversias en torno a cómo debería definirse el impacto (ver *United States General Accounting Office*, junio de 1997 (a)). En algunos casos, el impacto puede ocurrir sólo muchos años más tarde, de modo que ha existido la necesidad de desarrollar medidas provisorias o alternativas. Finalmente, aún cuando los impactos puedan ser especificados y medidos, existe una dificultad para especificar un estándar razonable para el desempeño correspondiente.

Estos problemas se han exacerbado cuando las medidas del impacto del desempeño han sido vinculadas con la planificación estratégica. Ha sido difícil establecer objetivos claros allí donde las misiones y las metas de la agencia no son claras. La clarificación de las mismas involucra temas de políticas y de “*trade-offs*”, que han sido deliberadamente mantenidos en términos poco concretos en el pasado por razones de orden político. Allí donde ha sido posible establecer metas estratégicas de largo plazo, ha sido difícil traducirlas en metas de desempeño anual, y los procesos han conllevado con frecuencia desacuerdos en torno a las prioridades, los enfoques y los roles.

De una importancia fundamental en lo que concierne a toda la cuestión de la evaluación de impactos, es saber si es posible establecer relaciones entre las actividades gubernamentales y los resultados. ¿En qué medida un programa produce resultados, o ha sido algún evento o tendencia externa el o la responsable de los mismos? Los impactos no siempre se pueden hacer corresponder con facilidad a los esfuerzos de la agencia, como por ejemplo en los programas reguladores o en la investigación científica. El desarrollo de medidas de desempeño puede requerir el análisis de los impactos de los programas.

Han existido también problemas relacionados con la calidad de la información y con cuestiones referidas a su adecuación y a su oportunidad. En particular, cuando la información es recolectada por terceras partes, tales como unidades gubernamentales subnacionales u organizaciones sin fines de lucro, puede ser difícil la verificación de la adecuación de la misma. Se han experimentado dificultades para establecer la línea de base a partir de la cual medir los programas realizados en los últimos años.

Finalmente, las medidas de desempeño tienen un elevado perfil político. La evaluación de programas afecta la sensibilidad política. Los programas que sean mostrados públicamente como fallando, de acuerdo con los criterios escogidos, están sujetos a recibir ataques. Los defensores de la evaluación sostendrían que así debe ser. Sin embargo, los programas no son entidades abstractas: son personas — administradores, políticos, clientes. Todos estos participantes de un programa, que tampoco son neutrales, pueden con justicia cuestionar las medidas de desempeño escogidas para evaluar el programa. Inclusive pueden poner en tela de juicio la calidad de la información y de su interpretación, dado que raramente los resultados de una evaluación dejan de lado toda ambigüedad. Por otra parte, los participantes astutos se asegurarán de que las medidas de desempeño muestren a sus programas desde una perspectiva favorable para justificar su prosecución. Los evaluadores pueden percibir sus medidas y criterios como neutrales y objetivos: los participantes los ven como un contrincante en la lucha política. La investigación de la evaluación no es infalible: sus resultados dependen de las medidas escogidas, de la calidad de los datos y de la interpretación de los hallazgos. Estudios concernientes a áreas complejas, tales como la prevención del delito, o los costos de la inmigración, pueden arribar a conclusiones bastante diferentes, y convertirse en armas para el debate político en curso.

Estos obstáculos parecen gigantescos, y algunos críticos han cuestionado el valor de la empresa en su totalidad. Aaron Wildavsky, por ejemplo, confesó: “Yo empecé pensando que era malo para las organizaciones no evaluarlas y terminé preguntándome por qué tienen que hacerlo” (Wildavsky, 1987: 458). La pregunta puede también ser formulada acerca de si es justificable, realista o ético para los denominados expertos en países ricos, con administraciones bien establecidas, pedir a sus pares en los países más pobres, que trabajan en un entorno con severas restricciones, que lleven a cabo reformas teóricas que ellos, con todas sus ventajas, encuentran extremadamente difíciles y aún imposibles. Este es un campo por el cual observadores y analistas de la administración pública han viajado antes — un paisaje contaminado por esquemas fallidos, del cual pocos, con la excepción de algunos consultores foráneos, se han beneficiado.

Contexto institucional

Aún no se ha hecho referencia al entorno en el cual debe ser conducida la medición del desempeño y la evaluación de programas. Ni siquiera en los países occidentales industrializados, con un servicio civil desarrollado y complejos vínculos entre los sectores público, privado y sin fines de lucro, las medidas del desempeño y la evaluación de programas

públicos han dejado de ser problemáticas. Verdaderamente, ha habido éxitos, pero los logros no han sido consistentes y han requerido un gran monto de esfuerzo. En los países pobres, ya estén clasificados como “en desarrollo” o “en transición” la introducción de las medidas del desempeño y de la evaluación de programas probablemente sea más dificultosa y confronte mayores problemas. Ni existe un sólo esquema del que pueda decirse que habrá de ser transferible a entornos tan radicalmente distintos. Ni es posible asegurar que todos, o si quiera algunos, de los gobiernos, necesariamente habrán de dar la bienvenida a un sistema de evaluación de sus agencias. Dado que las mediciones de desempeño y la evaluación de programas intrínsecamente presagian cambios, muchos habrán de ver estos cambios como amenazas, por lo que es altamente probable que surjan resistencias. Lamentablemente, se requiere capacidad para producir capacidad — información, experticia, recursos, y co-operación — recursos todos ellos que con frecuencia se muestran escasos.

Si la medición y la evaluación del desempeño son vistas simplemente como técnicas, ellas no pasarán de ser sólo una moneda de cambio de los consultores, quienes venderán sus sistemas, y dejarán tras ellos las cosas en el estado en que estaban, o un legado de cinismo acerca de la posibilidad de mejoramiento. Si tiene que haber algún grado de éxito, la evaluación y la medición del desempeño necesitan ser percibidas como un componente de desarrollo de la capacidad de gobierno. Esta concepción implicaría cambio en el sentido de una mayor transparencia, más información al público, más criticismo potencial de las agencias gubernamentales y una sociedad más abierta.

A medida que surgen nuevas ideas acerca de las reformas administrativas, se aprecia la tendencia por parte de sus proponentes, a recomendar su aplicación en todo lugar y en idénticas formas. La experiencia del pasado sugiere que no existe un sólo sistema apropiado para todos los lugares y todos los tiempos. El punto de vista convencional de ayer, hoy a menudo es percibido como ridículo. Más bien el desafío está en discernir los elementos esenciales que sustentan la recomendación actual, y adaptar lo que pueda parecer de utilidad a las circunstancias del entorno.

Los países varían en el tamaño, la economía, la cultura, las políticas y las instituciones. En algunos, la responsabilidad gubernamental no se encuentra bien establecida: los presupuestos son embrionarios y ocultan más de lo que revelan, en tanto que las instituciones de auditoría independientes están todavía por crearse. Mientras que algunos países están bien organizados y sus economías progresan, otros están en un estado de desorganización, al punto que los pagos gubernamentales no se hacen o se postergan, y la corrupción es sistémica.

La introducción de la evaluación y de medidas de desempeño en esta situación puede ser vista como una contribución crítica a la construcción de un servicio público profesional y al desarrollo de instituciones de gobierno viables. O puede ser percibida como una carga extra para un personal ya sobrecargado, y una distracción de los problemas más urgentes. Puede ser vista como un requerimiento esencial para la edificación de un sector público efectivo y democrático, o como un desarrollo que debería tener lugar más tarde, luego de haberse montado las estructuras básicas. Inevitablemente, la introducción de medidas del desempeño y de la evaluación de programas deberá ser justipreciada en términos políticos, de acuerdo a cómo ella afecta el balance de poder entre el ejecutivo y el legislativo, entre los niveles de gobierno y entre los sectores de la población. De ellas puede también esperarse que incidan en las demandas dirigidas a los gobiernos en relación con algunos programas, y en los esfuerzos por recortar o eliminar otros programas.

A pesar todo, el propósito central de la evaluación y de la medición del desempeño permanece: con una ejecución adecuada, constituye un instrumento de vigilancia del gobierno, un mecanismo para asegurar el valor del dinero público, una indagación esencial sobre la efectividad de los programas y de su contribución al bien público y un aporte para fortalecer la confianza en los gobiernos. La publicación de las medidas del desempeño contribuye con la difusión de la información acerca de los gobiernos, y con la creación y el mantenimiento de una ciudadanía informada y una sociedad civil motivada. Igualmente, las medidas del desempeño sustentan y fortalecen el profesionalismo en el servicio público, en contra de la idea de la satisfacción de intereses personales por parte de los burócratas y los políticos corruptos. Pero para que las medidas de desempeño hayan de ser tomadas seriamente y para que sean de valor, mucho depende de cómo son introducidas e implementadas en un entorno complejo, turbulento y stressado.

Cada país tiene que determinar sus propias necesidades y usos de la medición del desempeño y de la evaluación, y también considerar la mejor forma de implementarlos en sus propias instituciones políticas y administrativas. Cualquiera sea la dirección que se escoja, será necesario asignar responsabilidades para asegurar que las reformas son realmente emprendidas en una forma significativa. Al igual que todas las innovaciones, el compromiso desde las instancias superiores es esencial, como lo es la comprensión por parte de los administradores en la base de la organización. Con demasiada facilidad la medición del desempeño puede tornarse un simple ejercicio de auto-justificación, y por lo tanto desestimado o no tomado seriamente consideración. Por consiguiente, las medidas del desempeño y los estudios evaluativos no deberían concebirse en forma aislada, sino como resultado la consulta entre funcionarios del ejecutivo y del legislativo,

y entre los varios interesados que las desarrollarán y harán uso de ellas. El propósito es lograr el consenso y la cooperación en el interés por alcanzar una verdadera comprensión de cómo operan los programas y de cómo podrían operar mejor.

Existe una variedad de opciones para diseñar instituciones que aseguren la calidad en la medición del desempeño y en la evaluación. Los arreglos institucionales podrían incluir, por ejemplo, una unidad especial agregada al despacho del Primer Ministro, oficinas de evaluación en departamentos más grandes o grupos de agencias, extensiones de las funciones de las oficinas de auditoría responsables ante la legislatura y/o el ejecutivo, funciones adicionales para auditores en general y para auditores internos, participación de los ciudadanos, comisiones, “*think tanks*” independientes, entre otros. La información podría difundirse, por ejemplo, a través de documentos presupuestarios, informes financieros, informes regulares sobre indicadores en áreas particulares o que afectan a grupos de población específicos, informes especiales solicitados por los legisladores, informes de auditoría, y por los medios de comunicación. Indudablemente que surgirán objeciones en relación con el hecho de que tal información podría ser utilizada con fines políticos, o para fortalecer posiciones grupales, pero la democracia exige información y conocimiento sobre el trabajo de sus instituciones, y cuanto mejor sea la calidad de dicha información y cuanto más libremente fluya, menos probabilidades habrá de que ella sea manipulada o distorsionada.

Finalmente, si tiene que producirse y utilizarse una información de calidad sobre el desempeño, es necesario desarrollar experticia y capacidad de juicio profesional entre los responsables de su producción y utilización. No es necesario inventar la pólvora, desde el momento en que se cuenta con una experiencia considerable en materia de desarrollo e implementación. Ya tienen a disposición manuales y ejemplos que pueden adaptarse a usos específicos, y el tema puede agregarse fácilmente a los currícula de entrenamiento ya existentes.

El balance final

Así como se espera que las organizaciones privadas informen anualmente a sus accionistas, clientes y público en general acerca de su desempeño, igualmente debería esperarse que las organizaciones públicas hicieran lo mismo, no sólo en términos de adecuación financiera, sino también en términos de progreso sustantivo, competencia profesional y responsabilidades públicas. Durante un tiempo demasiado largo esta tarea ha sido desestimada, pero en las últimas décadas se han realizado avances en la medición cuantitativa de las actividades del gobierno. Lo que se

necesita es hacer énfasis en la evaluación cualitativa de los programas del sector público.

Existen buenas razones para esperar que todas las agencias gubernamentales produzcan informes anuales. Estos documentos deberían ser informativos, explicando al público por qué existe la agencia, qué hace y qué ha estado haciendo.

En primer lugar, los informes deberían ser educativos por naturaleza, informando al público acerca de lo que debería conocer y contribuyendo al desarrollo de una sociedad civil vibrante y activa. Las declaraciones anuales actuarían fortaleciendo la confianza en las agencias públicas, incentivando la moral tanto de los servidores públicos como de los clientes, llenando vacíos en las relaciones públicas para contrarrestar el criticismo mal informado, y mostrando qué bienes y servicios públicos suministra el gobierno.

En segundo lugar, tales informes son también necesarios con fines de gerencia interna, contribuyendo con una efectiva supervisión, con la toma de decisiones y con la auditoría. Ellos suministran la información de la línea de base para la toma de decisiones y la revisión retrospectiva del desempeño.

En tercer lugar, dados los grandes montos de dinero gastado por los gobiernos modernos, parecería razonable utilizar al menos una pequeña parte del mismo para determinar cuán bien o mal ha sido gastado y qué se ha conseguido con ello. Desde este punto de vista, los gastos para la evaluación y la medición del desempeño deberían verse como inversiones más que como meras erogaciones superfluas.

En este enfoque existen tanto problemas como riesgos. La producción de publicaciones informativas regulares requiere de recursos que faltan en muchos países. Existen muy pocas personas disponibles, que tengan capacidad para comprender e implementar la asunción de responsabilidades y la transparencia. La información, fuera de los países occidentales industrializados, es escasamente difundida, aún en lo que respecta a problemas básicos, y aún más en áreas complejas y ambiguas como son las medidas cualitativas del desempeño del gobierno, tales como la justicia o humanitarias. La innovación y la creatividad son necesarias para procesar datos útiles, informativos y confiables, y presentarlos regularmente en una forma comprensible. Obviamente, se corren riesgos en esta empresa: los datos pueden ser manipulados, pueden ser mal representados, o pueden desviarse con fines de propaganda partidista. La información honesta acerca del desempeño puede también ser mal utilizada para desprestigiar o destruir las organizaciones gubernamentales. Puede también constituirse en una amenaza, un medio para justificar el recorte de fondos para agencias que no gozan de popularidad, para victimizar a ciertos individuos o para degradar las condiciones del empleo público.

Pero los problemas pueden superarse, y los riesgos pueden disminuirse mediante una planificación y una implementación cuidadosas. Existe una considerable experticia a disposición para ser aprovechada en las áreas de evaluación, medición del desempeño y recolección de datos, así como en la elaboración de documentos informativos. Esta es un área en la cual la difusión de la información a través de canales de *Internet* podría ser factible y relativamente barata. Es también un área que se presta para la adaptación flexible, dado que una amplia variedad de enfoques han sido actualmente bien documentados. La experiencia ha mostrado que las medidas del desempeño y la evaluación deberían ser adecuadas a la tarea, y que la aceptabilidad y la implementación se ven facilitadas allí donde aquellos afectados por las medidas y aquellos que podrían hacer uso de ellas se encuentran activamente involucrados en el diseño de las mismas. No hay un único esquema: la asistencia es necesaria para proveer recursos y orientación. Inclusive pequeños pasos en esta dirección pueden contribuir a aumentar la transparencia y a aproximarse a una sociedad más democrática, aspectos que la comunidad internacional ha apoyado con entusiasmo en la Resolución de Naciones Unidas 50/225, cuyas partes más relevantes llaman a un mejoramiento de la efectividad de las actividades de la ONU en la administración pública el desarrollo (Artículo 11) a través del procesamiento y del acceso a la información, promoviendo el entrenamiento y la investigación, la defensa y el intercambio de experiencias, servicios de asesoría, asistencia técnica, formación de capacidades y desarrollo de los recursos humanos (Artículo 13).

En esta perspectiva, el 50º aniversario de los Programas de Administración Pública y Finanzas de las Naciones Unidas, debería no sólo reflejar los progresos alcanzados, sino también ofrecer lineamientos acerca de dónde centralizar los esfuerzos en el futuro. Este documento adelanta cuatro grandes recomendaciones.

(1) El seguimiento, la medición y la evaluación del desempeño del sector público deberían seguir siendo una prioridad para la investigación y la formación de capacidades. Sin tales índices, pocos gobiernos conocen cuán bien (o cuán mal) se están desempeñando, cuánto progreso (o atraso) están impulsando y cómo podrían hacer lo mejor para ser capaces de satisfacer las expectativas del público. Existen indicadores claves acerca de dónde los gobiernos deberían centrarse para elevar más su desempeño en el corto plazo, para planificar los esfuerzos del futuro, para mejorar rápidamente el desempeño general y mediante inversiones relativamente modestas. Sin una prosecución del desarrollo de estos indicadores o *benchmarks*, los gobiernos seguirán tal vez enfrentando un criticismo injustificado de su desempeño, no podrán saber realmente dónde están y hacia dónde se orientan, dónde se localizan sus áreas débiles y dónde sería preciso priorizar las inversiones. Los gobiernos seguirán

perpetuando programas que no tienen valor, que son dispendiosos o que brindan pocos beneficios, y que dan lugar a grandes insatisfacciones y protestas del público.

(2) La comunidad internacional no debería dudar en cuanto a **promover y aplicar medidas del desempeño a las organizaciones internacionales**, incluyendo a la familia de organizaciones de la ONU. Tal como se recomienda en este documento, todas las organizaciones internacionales deberían producir informes anuales contentivos de informaciones similares a las de aquellos emitidos por las grandes corporaciones multinacionales, incluyendo medidas del desempeño. Sin *benchmarks*, los estados miembros no pueden juzgar en forma apropiada a estas organizaciones internacionales y sus progresos, no pueden evaluar la validez y el valor del progreso de la humanidad, no pueden identificar dónde el progreso ha sido débil y no pueden determinar dónde inversiones relativamente bajas pueden producir los cambios más rápidos y más beneficiosos para ella.

(3) Las medidas detalladas, que toman mucho tiempo y esfuerzos para aplicarlas, no son de tan alta prioridad como lo son las **medidas relativamente sencillas, claras y directas, que muestran los resultados o los productos de un programa**, por ejemplo cuántas vidas de niños se han salvado, cuántas personas han sido provistas de alimentos y abrigo, cuántos trabajadores han encontrado empleo, etc. Se ha prestado una atención excesiva a la legalidad, a la probidad financiera y a la magnificación burocrática (cuántas personas están empleadas, cuánto espacio ocupan las oficinas, cuán abultado es el presupuesto), por más que éstos sean importantes pre-requisitos para futuros progresos; demasiado poca atención se ha dedicado a los resultados, a los productos, a los impactos y a los beneficios para la gente que depende del desempeño del sector público. Los gobiernos tienen que mostrar resultados que la gente común entienda y aprecie, aún si los resultados son desalentadores o revelan que los programas han sido ineficaces o inadecuados, y aún cuando sus fondos hayan sido desviados para usos impropios, cosa que de todos modos el público debería conocer.

(4) Las implicaciones son claras en términos del **alto nivel de compromiso**, tanto internacional como nacional, para invertir en el área de seguimiento, medición y evaluación del desempeño del sector público. Ha habido demasiadas conversaciones y poca acción. Se han realizado demasiados estudios exploratorios y demasiado poca implementación. Se ha enfatizado demasiado en lo periférico y demasiado poco en los problemas centrales. Demasiado poca inversión se ha realizado, muy pocas personas se han involucrado en esta área de la administración y de las finanzas públicas, y disponen de demasiado pocos recursos para trabajar en ella. Esta área, en su totalidad, se encuentra subdesarrollada, si se considera

que los gobiernos y los programas del sector público han existido al menos desde hace 6.000 años de historia escrita. Hay demasiado poco que mostrar hacia fines del siglo XX.

Apuntes

- * Artículo publicado en la Revista del CLAD Reforma y Democracia, Caracas, nº 12, (oct. 1998).

Bibliografía

- ORGANIZATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT. (1996), *Ministerial Symposium on the Future of Public Services*. Paris: OCDE, march.
- ROSEN, Ellen D. (1993), *Improving Public Sector Productivity: concepts and practice*. Newsbury Park Ca: Sage Publicantions.
- SWISS, James E. (1991), *Public Management Systems: monitoring and managing government performance*. Englewood Cliff's N.J.. Prentice Hall.
- UNITED KINGDOM NATIONAL AUDIT OFFICE. (1992), *Performance Measurement in the Public Sector: draft discussion paper prepared for the World Bank*. London: HMSO, june.
- UNITED STATES GENERAL ACCOUNTING OFFICE. (1992), *Program Performance Measures: federal agency collection and use of performance data*. Washington D.C. (GAO/GGD - 92-65).
- _____. (1993), *Performance Budgeting: state experiences and implications for the Federal Government*. Washington D.C. (GAO/AFMD - 93-41), february.
- _____. (1995), *Managing for results: experiences abroad suggest insights for Federal Management reforms*. Washington D.C. (GAO/GGD - 95-120), may.
- _____. (1997), *Performance Budgeting: past initiatives offer insights for GPRA implementation*. Washington D.C. (GAO/AIMD - 97-46), march.
- _____. (1997), *Managing for results: analytic challenges in measuring performance*. Washington D.C. (GAO/HEHS/GGD - 97-138), may.
- _____. (1997), *The Government Performance and Results Act: 1997 government wide implementation will be uneven*. Washington D.C. (GAO/GGD - 97-109), june.
- _____. (1997a), *Managing for Results: regulatory agencies identified significant barriers to focusing on results*. Washington D.C. (GAO/GGD - 97-38), june.
- WILDAVSKY, Aaron. (1987), "The Self Evaluation Organization" en Jay M. Schafritz and Albert hyde, "Classics of Public Administration", 2nd edition, Pacific Grove. California: Brooks Cole.

Enfoques e orientações para o monitoramento, a mensuração e a avaliação de desempenho em programas do setor público.

Gerald E. Caiden e Naomi J. Caiden

O artigo aborda a evolução da mensuração do desempenho e da avaliação de programas nas organizações do setor público, e destaca-se a significação da mesma na atualidade, fazendo alusão a casos de países em desenvolvimento. Comenta os principais aspectos da implementação de um sistema de avaliação e mensuração de desempenho, e discute seus alcances e limitações.

A busca de medidas apropriadas para a avaliação quantitativa e qualitativa das organizações do setor público não é nova. Sem dúvida, nos últimos anos, a preocupação em torno da avaliação dos programas tem aumentado, ao mesmo tempo em que as reformas do Estado avançam com novos critérios. Assim, a avaliação e a mensuração do desempenho passaram a ser parte destas reformas, orientadas para maior flexibilização da gerência, *accountability*, descentralização, privatização e foco nos usuários.

A avaliação e a mensuração do desempenho implicam na obtenção e interpretação de dados sobre vários aspectos do processo produtivo: insumos, produtos, resultados, impacto, produtividade e satisfação do usuário, entre outros. Estes aspectos distintos põem em evidência a complexidade da tarefa e as dificuldades inerentes a mesma. Assim, por exemplo, enquanto é relativamente simples a obtenção de dados sobre insumos, esta pode ser extremamente complexa quando se trata de avaliar o impacto de um programa, onde a dimensão qualitativa é mais relevante que a quantitativa, e as metas não admitem uma mensuração precisa.

A introdução da avaliação e da mensuração do desempenho possivelmente encontra ainda mais dificuldades nos países pobres do que nos ricos. É possível que nem todos os governos estejam de acordo em realizar uma avaliação do funcionamento de seus programas por diferentes razões. Por outro lado, para alguns, esta tarefa poderia ser postergada para o momento em que as estruturas institucionais estivessem consolidadas e que o pessoal tivesse capacitação mais apropriada para sua realização. Conseqüentemente, cada país terá que determinar não só a oportunidade para implementar a mensuração do desempenho, como também escolher que tipo de medidas procurará desenvolver, de acordo com as características próprias do contexto em que elas serão aplicadas.

Conclui-se que a mensuração do desempenho constitui um instrumento que pode colaborar com uma maior eficiência e eficácia dos serviços públicos, e propiciar maior abertura da administração, com maior transparência e aumento da responsabilização por parte dos funcionários encarregados de tais serviços. Portanto, é recomendável a publicação periódica dos dados concernentes ao desempenho das organizações públicas, tal como o fazem as grandes corporações privadas.

Enfoques y lineamientos para el seguimiento, la medición y la evaluación del desempeño en programas del sector público

Gerald E. Caiden y Naomi J. Caiden

Se reseña la evolución de la medición del desempeño y de la evaluación de programas en las organizaciones del sector público, y se destaca la significación de la misma en la actualidad, haciendo alusión al caso de los países en desarrollo. Se comentan los principales aspectos de la implementación de un sistema de evaluación y medición del desempeño, y se discuten sus alcances y sus limitaciones.

Gerald E. Caiden es egresado de la Escuela de Economía y Ciencias Políticas de Londres, consultor, investigador y administrador de organizaciones internacionales como Banco Mundial y ONU.

Naomi J. Caiden cursó estudios en la Escuela de Economía y Ciencias Políticas de Londres, actualmente es profesora de administración pública en la Universidad del Estado de California en Los Angeles.

La búsqueda de medidas apropiadas para la estimación cuantitativa y cualitativa de las organizaciones del sector público no es nueva. Sin embargo, en los últimos años la preocupación en torno a la evaluación de los programas se ha incrementado, al tiempo que se adelantan reformas del Estado con nuevos criterios. Así, la evaluación y la medición del desempeño han pasado a ser parte de estas reformas, orientadas hacia una mayor flexibilización de la gerencia, la asunción de nuevas responsabilidades por los servicios públicos, descentralización, privatización y atención a las necesidades de los usuarios.

La evaluación y medición del desempeño implica la obtención y la interpretación de datos sobre varios aspectos del proceso productivo, a saber: insumos, productos, resultados, impacto, productividad y satisfacción del usuario, entre otros. Estos distintos aspectos ponen en evidencia la complejidad de la tarea y las dificultades inherentes a la misma. Así, por ejemplo, mientras es relativamente sencilla la obtención de datos sobre insumos, ella pueda ser extremadamente compleja cuando se trata de justipreciar el impacto de un programa, donde en ocasiones la dimensión cualitativa es más relevante que la cuantitativa, y las metas no admiten una medición precisa.

La introducción de la evaluación y la medición del desempeño posiblemente encuentre aún más dificultades en los países pobres que en los países ricos. En efecto, es posible que no todos los gobiernos estén de acuerdo en realizar un escrutinio del funcionamiento de sus programas, por distintas razones. Por otra parte, para algunos esta tarea podría postergarse para el momento en que las estructuras institucionales estén consolidadas y que el personal tenga una capacitación más apropiada para su realización. En consecuencia, cada país tendrá que determinar no sólo la oportunidad para implementar la medición del desempeño, sino también escoger qué tipo de medidas procurará desarrollar, acorde con las características propias del entorno en que ellas serán aplicadas.

Se concluye que la medición del desempeño constituye un instrumento que puede colaborar francamente con una mayor eficiencia y eficacia de los servicios públicos, y propiciar una mayor apertura de la administración, con mayor transparencia y asunción de responsabilidades por parte de los funcionarios encargados de dichos servicios. Por tanto, es recomendable la publicación periódica de los datos concernientes al desempeño de las organizaciones públicas, tal como lo hacen las grandes corporaciones privadas. Esta práctica es beneficiosa para la democracia, aún si los datos son negativos para la organización, y se corra el riesgo de que sean manipulados, debiéndose recordar que a mayor transparencia en los procedimientos, menos probabilidades habrá de que los datos sean objeto de manejos dolosos.

Approaches and guidance for monitoring, measuring and assessing performance in public-sector programmes

Gerald E. Caiden and Naomi J. Caiden

The article approaches the evolution of performance measurement and programme assessment in public-sector organisations, and it highlights its meaningfulness in present times, referring to cases of developing countries. It elaborates on the main aspects of the implementation of an assessment and measurement system and discusses its scope and limitations.

The quest for appropriate measures for the qualitative and quantitative assessment of public-sector organisations is not new. Doubtlessly, in recent years the concern about programme assessment has increased, at the same time, as state reforms advance with new criteria. Thus, performance assessment and measurement have become part of those reforms, geared to a more flexible management, accountability, decentralisation, privatisation and focus on users.

Performance assessment and measurement imply obtaining and construing data on several aspects of the productive process: inputs, outputs, outcomes, impact, productivity and user's satisfaction, *inter alia*. These distinct aspects expose the complexity of the task and its inherent difficulties. Thus, for example, while it is relatively easy to gather data on inputs, data collection can be highly complex when it is about assessing the impact of a

programme, in which the qualitative dimension is more relevant than the quantitative aspect and targets do not allow the performance of an accurate measurement.

The introduction of performance assessment and measurement is likely to meet even more difficulties in poor countries than in affluent ones. It is possible that not all governments agree to carry out an assessment of the operation of their programmes for different reasons. On the other hand, for some governments, this task might be postponed to the moment when institutional structures are consolidated and the personnel are more appropriately trained for carrying it out. As a consequence, each country will have to determine not only the opportunity for implementing performance measurement, but also choose which type of measures will develop, according to the specific characteristics of the context in which they will be applied.

The conclusion is that performance measurement represents an instrument that may provide more efficiency and effectiveness to public services and propitiate more openness for administration, with more transparency and more accountability. It is therefore advisable to periodically publish the data regarding the performance of public organisations; such as large private corporations do.