

APÉNDICE B

EL ARTE DE MEDIR LOS RENDIMIENTOS

El gobierno es famoso por sus interminables cifras y formas. Para un extraño, parece una industria que preste mucha atención a los números. La gente del gobierno siempre está contando algo o agitando algún informe estadístico. Pero el grueso de su contabilidad se centra en los datos: cuánto se gasta, a cuánta gente se sirve, qué servicio recibe cada persona. Muy raras veces se concentra en las entradas, en los *resultados*.

En parte es así porque medir los resultados es muy difícil. El medir los resultados en la empresa es muy sencillo. Medir los resultados en el gobierno, no lo es. Normalmente lleva varios años desarrollar medidas adecuadas: el primer intento de una agencia a menudo por desgracia fracasa. Puede que mida sólo datos, no entradas. Puede definir los datos de un modo demasiado estrecho, haciendo que los empleados se concentren únicamente en unos pocos resultados que la organización desea realmente cumplir. Puede desarrollar tantas medidas que los empleados no puedan decir en qué hay que concentrarse.

Incluso Sunnyvale, en California, quizá el gobierno más sofisticado que conocemos, cuando se trata de medir rendimientos, sigue utilizando medidas que piden reajustes. Como Sunnyvale, Phoenix, cientos de otras organizaciones gubernativas han desarrollado sus propios sistemas de medida, y por ello, han aprendido un número de lecciones básicas:

- 1. Existe una gran diferencia entre medir procesos y medir resultados.** Cuando las organizaciones públicas se ponen a me-

dir rendimientos, sus gestores normalmente diseñan listas que miden lo bien que llevan algunos procesos administrativos: a cuánta gente sirven; con qué velocidad les sirven; qué porcentaje de solicitudes se llenan en un período de tiempo dado. Básicamente, miden su volumen de datos. Pero los *datos* no garantizan *entradas*. Una escuela vocacional puede ir poniendo en circulación cada vez más graduados de un programa de soldadura, por ejemplo. Pero si aquellos que se gradúan no pueden encontrar trabajo de soldador, ¿para qué sirve el programa? Tal vez esté generando datos impresionantes sin generar ninguna entrada positiva.

La tendencia a concentrarse en los procesos es natural: los gestores miden lo que hacen sus agencias, y en las organizaciones regidas por normativas, la gente considera que sus trabajos consisten en seguir ciertos procesos bajo ciertas normativas. Si siguen dichos procesos fielmente y producen el esperado volumen de datos, están haciendo su trabajo. Pocas veces piensan en las entradas: el impacto que las actividades tienen en aquellos que la agencia debe servir. Sin embargo un proceso ejecutado a la perfección, si fracasa a la hora de conseguir las entradas deseadas, es una pérdida de tiempo y dinero. El Centro Nacional de Tribunales del Estado¹ proporciona un buen ejemplo. Junto al Departamento de Justicia de EE.UU., intentó crear unas normas para las actuaciones de los tribunales. El propósito, expresado de un modo explícito, era concentrarse no «en las estructuras y maquinaria de los tribunales» sino «en sus rendimientos (*aquello que realmente conseguían los tribunales con los medios puestos a su disposición*)» (énfasis añadido). Pero el grupo se vio luchando continuamente con lo que el doctor Ingo Keilitz, que dirigía el proyecto, llama: «proceso reptante». Algunos miembros suponían que «la buena gestión es un resultado», recuerda Keilitz. «Y contestábamos: “Pero la buena gestión no es un fin en sí misma”. Discutíamos hasta acalorarnos.»

Evidentemente el medir procesos puede ser de utilidad. La

1. *Trial Court Performance Standards with Commentary* (Williamsburg, Va.: National Center for State Courts, 1990), pág. 1.

buena gestión es importante, y la medición de procesos puede ayudar a las organizaciones a ver cómo pueden mejorar su gestión. Las organizaciones que utilizan, por ejemplo, el Total Quality Management de Deming, miden constantemente sus procesos internos de modo que pueden ver dónde están los problemas y corregirlos. Además, los resultados deseados por cualquier organización son a veces difíciles de medir, o no se hacen evidentes hasta que pasa un tiempo. En estos casos, las organizaciones a veces eligen mediciones de procesos que parecen ser aproximaciones fiables a los resultados definitivos. Por ejemplo, la Asociación Ben Franklin de Pennsylvania mide cosas como la cantidad de inversión privada atraída para igualar cada una de sus subvenciones, suponiendo que los proyectos que atraigan una significativa cantidad de capital tienen más posibilidades de contribuir al crecimiento económico que los proyectos que tienen dificultades a la hora de atraer capital privado.

El problema surge cuando las organizaciones *sólo* miden los procesos, como tantas suelen hacer. El Fox Valley Technical College mide muchos tipos de procesos: el número de cursos programados, pero luego abandonados; el uso de las técnicas de evaluación que acentúan la destreza de la competencia más que la capacidad de hacer exámenes escritos; la cantidad de instrucción basada en ordenadores que ofrece; etc. Pero si no mide también cuántos licenciados consiguen trabajos en los campos para los que han sido formados, su satisfacción y la satisfacción de los que los emplean, puede crear mejores cursos que tengan como resultado cada vez menos puestos de trabajo. Una vez que los robots han sustituido a los soldadores, no tiene sentido trabajar para mejorar nuestros cursos de soldador.

2. Existe una gran diferencia entre medir la eficacia y medir la efectividad. La eficacia es una medida de lo que cuesta cada unidad de resultado. La efectividad es una medida de la calidad de este resultado: ¿Con que precisión consigue el resultado deseado? Cuando medimos la eficacia, sabemos lo que cuesta conseguir un resultado específico. Cuando medimos la efectividad, sabemos si nuestra inversión tiene valor. No hay nada

más ridículo que hacer de un modo más eficaz algo que debe dejar de hacerse.

Tanto la eficacia como la efectividad son importantes. Pero cuando las organizaciones públicas empiezan a medir sus rendimientos, a veces miden sólo su eficacia. Normalmente, un Departamento de Defensa tradicional mide lo que cuesta alojar y alimentar a sus tropas; y se esfuerza constantemente por rebajar la cifra. Pero como Bob Stone explica en su documento de cuatro páginas «Criterios de Interpretación del Departamento de Defensa»: El fin *no* es minimizar el coste del ciclo de vida de los servicios, sino maximizar la actuación de la gente que utiliza dichos servicios».

El público evidentemente quiere un gobierno eficaz, pero aún más un gobierno efectivo. Los ciudadanos tal vez se complazcan en gastar menos en educación por alumno que el resto de Estados, pero si sus escuelas son las peores del país, no tienen muchas razones para estar felices por mucho tiempo. Pueden disfrutar de unos impuestos bajos, pero si ello significa que pasan una hora en autopistas embotelladas para llegar al trabajo, normalmente votarán para invertir en un sistema de transportes más efectivo.

El centrarse en la eficacia más que en la efectividad también tiende a alienar a los funcionarios. Cuando los gobiernos acenúan el coste de cada unidad de trabajo, a menudo desarrollan una mentalidad de antifaz que menosprecia la inteligencia y destreza de sus trabajadores. La mayoría de los empleados quieren ser efectivos. La mayoría gustosamente haría lo que fuera necesario para aumentar el impacto de su organización. Pero si sus superiores únicamente se concentran en su eficacia —en lo rápido que hacen cada unidad de trabajo— empezarán a sentirse como si estuvieran en una cadena de montaje.

George Britton, un asistente de gestor en Phoenix, señala que su ciudad se inclinó mucho por medir la eficacia en la década de los 70. «Consideramos que cada barrendero tenía que cubrir 104 km de suelo por día»; y afirma:

Pero este principio se vino abajo, puesto que cayó en una filosofía de ojos vendados. Se perdió de vista la pregunta: ¿para qué hace-

mos este servicio? Los gobiernos hacen muchos servicios de un modo muy eficaz, sin preguntarse porqué los hacen. Hemos perdido de vista la relación entre eficacia y efectividad. Sólo eres eficaz si haces algo que se necesite hacer.

3. Existe una gran diferencia entre los «resultados de programa» y los más amplios «resultados políticos». Cuando el centro Nacional de Tribunales del Estado desarrolló sus estándares de actuación, presentó estándares sobre resultados que medían la satisfacción pública por la accesibilidad, objetividad, fiabilidad, velocidad, independencia política, y responsabilidad de un tribunal. Para medirlo los tribunales utilizaron encuestas públicas, grupos seleccionados, etc. Pero ninguna de estas medidas afrontaba los resultados más amplios y de mayor interés para el público: la tasa de crímenes, la seguridad pública, la tasa de convictos, la tasa de reincidentes, la justicia para las víctimas de delitos, o la satisfacción por el modo en que se resolvían los juicios. Los jueces, evidentemente, dirán que los tribunales no son las únicas instituciones responsables por cosas como la tasa de delitos, las tasas de convictos, y las percepciones públicas de la seguridad. Estarán en lo cierto. Pero se trata de los resultados por los que más se preocupa la gente.

El dilema se presenta a muchas organizaciones públicas. Pueden medir los resultados de sus programas o actividades específicas, pero dichos números no son tan importantes como ciertas medidas más amplias. Un programa de trabajo destinado a bienestar social debe medir su tasa de colocación y los salarios de los colocados, por ejemplo. Pero ¿qué sucede con la cantidad de gente que se apunta al bienestar social, el tiempo que está, y la medida de la carga? Los gestores del programa de trabajos pueden decir que ellos son responsables del número de colocados, pero no por el conjunto de personas que acude a bienestar social.

De nuevo, estarán en lo cierto. Los últimos números están afectados por cuestiones políticas más amplias: a quién se selecciona, qué grado de atracción tiene el bienestar social comparado con trabajos no cualificados y de sueldo bajo; y cuánto trabajo de sueldo bajo hay disponible. Pero si el departamento

de bienestar social no mide ambos conjuntos de números, puede creer que está haciendo una gran labor sacando a la gente del bienestar social, mientras sus listas realmente están aumentando. Esto es precisamente lo que sucedió a muchos Estados en la década pasada. Massachusetts' Employment and Training (E.T.) Choices Program (el Programa de Elección de Empleo y Formación de Massachusetts) está considerado, con justicia, uno de los mejores de la nación. Pero aunque colocó aproximadamente 10.000 personas al año² entre 1983 y 1990, los beneficiarios del departamento AFDC empezaron siendo 90.000, nunca bajaron de 84.000, y pasaron de los 100.000 cuando dio comienzo la recesión.

Estadísticas de este tipo descubren la importancia de medir tanto los *resultados de los programas como los resultados políticos*. Programas como Elecciones E.T. precisan información de lo bien que están consiguiendo sus metas. Los gestores de Elecciones E.T. recogen dicha información religiosamente y la utilizan continuamente para ajustar y mejorar sus esfuerzos. Pero también precisan de información de aspectos más amplios, como el porcentaje de población que está en la asistencia social y en la pobreza. Con dicha información, se hace evidente que Elecciones E.T. está innovando dentro de un más amplio sistema de asistencia social que tiene poderosos incentivos para que la gente entre y esté en ella. Podemos esperar que en algún momento dicha información desencadene esfuerzos para cambiar el sistema más amplio, en lugar de simplemente seguir poniendo el acento en emplear a gente.

El cuadro de las páginas 475-476, esboza la diferencia entre datos (procesos) y entradas (resultados), entre resultados de programa y resultados políticos, y entre eficacia y efectividad. Luego ilustra dichas diferencias para varios servicios públicos distintos. Por utilizar como ejemplo el barrer una calle, ilustra que debemos calibrar la *salida*, o proceso, midiendo el número de

2. David Osborne, *Laboratories of Democracy* (Boston: Harvard Business School Press, 1988) págs. 200-206; y Scott Lehigh y M. E. Malone, «Weld: Revenue Loss Threatens '92 Budget», *Boston Globe*, 23 de abril de 1991.

kilómetros barrido. Pero si deseamos medir la *entrada*, o el resultado, tenemos que medir la limpieza de la calle; tal vez mediante la valoración de observadores objetivos entrenados. Esta última medida se convertiría en el *resultado del programa*. Un *resultado político* contemplaría un tema más amplio: ¿cuánta basura dejan los ciudadanos en las calles, y qué efectividad tiene la política pública a la hora de hacer disminuir la cantidad?

Para determinar la *eficacia del programa*, una organización simplemente medirá el coste por kilómetro barrido. Pero para determinar la *eficacia política*, tendría que medir el coste de lograr un nivel deseado de limpieza de la calle, sea cual sea el método, barrer, prevención, o autoayuda de la comunidad. Por último, para medir la *efectividad del programa*, una ciudad debe medir la satisfacción del ciudadano con el grado de limpieza de la calle. Pero para medir la *efectividad política*, deberá preguntar a los ciudadanos si quieren que se gaste su dinero en limpiar las calles, o qué usos alternativos, como pueda ser la construcción o pavimentación, serán preferibles.

Para aquellos que estén interesados en saber más sobre las medidas de los rendimientos, un número de expertos han escrito ampliamente sobre el tema.³ Harry Hatry, del Instituto Urbano, ha publicado una serie de libros, que están concretamente destinados a gobiernos locales. Jack Brizius, Michael Campbell, y Roger Vaughan han escrito, para el Consejo de Asesores Políticos de los Gobernadores y la Corporación para el Desarrollo de la Empresa, sobre la medida de los rendimientos en los gobiernos de Estado. En 1990, La Governmental Accounting Standards Board (GASB) (Junta Gubernamental de Normativas Con-

3. Harry P. Hatry, Louis H. Blair, Donald M. Fisk, John H. Greiner, John R. Hall, Jr. y Philip S. Schaenman, *How Effective Are Your Community Services? Procedures for Monitoring the Effectiveness of Municipal Services* (Washington, D. C.: Urban Institute and International City Management Association, 1977); Jack A. Brizius y Michael D. Campbell, *Getting Results* (Washington DC.: Council of Governor's Policy Advisors, 1991); Richard E. Brown, Meredith C. Williams y Thomas P. Gallagher, *Auditing Performance in Government: Concepts and Cases* (Nueva York: Ronald Press, 1983); Harry P. Hatry et al. (comps.), *Service Efforts and Accomplishments Reporting: Its Time has Come: An Overview* (Norwalk, Conn.: Governmental Accounting Standards Board, 1990); y Roger Vaughan, «Is It Working», *Entrepreneurial Economy Review* (julio-agosto 1989): 3-7.

tables) publicó un informe de investigación sobre las valoraciones de rendimientos, llamado *Service Efforts and Accomplishments Reporting: Its Time Has Come*, que proporciona un panorama útil. GASB está desarrollando también detallados informes de investigación en varios campos específicos, como la educación, el desarrollo económico, los hospitales, el transporte público, y el tratamiento de aguas. (Para citas específicas véanse las notas de editor, pág 473.)

Los expertos tienen la tendencia a estar de acuerdo en una serie de lecciones:

Llevar a cabo tanto análisis cuantitativos como cualitativos. Algunos resultados valiosos son imposibles de cuantificar. Otros exigen tanto papeleo y gasto que es mejor no cuantificarlos. Otros son cuantificables, pero nadie puede asegurar qué programa en cuestión fue el responsable de su producción. Por todas estas razones, es importante combinar las medidas cuantitativas con la valoración cualitativa. Los buenos gestores pueden tener muchas aclaraciones sobre las actuaciones cuando observan los números correspondientes, pero también pueden tener valiosas aclaraciones si dedican su tiempo a observar el programa, la agencia o el proveedor, hablan con los trabajadores, y escuchan a los clientes.

Si toda la valoración de las actuaciones se hace mediante números, los proveedores de servicios tal vez aprendan también a jugar con los números. «Observamos los (de empleo e inscripción) números, pero no queremos que nuestros centros se obsesionen con ellos», explica Suzanne Teegarden, directora del Programa de Servicios Industriales de Massachusetts.

Nuestros peores centros son aquellos dirigidos por los números. Uno de ellos, en caso de que la meta sea inscribir a 250 personas, podemos dar por seguro que inscribirá a 252. No tiene mucha energía y no se arriesga. Nuestros mejores centros son aquellos que se cuidan de verdad, que son creativos, que primero actúan y después preguntan. Conocemos bien a los centros y por ello sabemos lo que significan los números.

¿QUÉ MEDIR?

	Definición general	Barrido de calle	Bienestar: formación del trabajo
Salida (o proceso)	Volumen de unidades producidas	Kilómetros barridos	Número de gente formada
Entrada (o resultado)	Calidad/efectividad de producción: grado en que crea los resultados deseados	Tasa de limpieza de la calle	Número de gente colocada en trabajos, trabajando y fuera de la asistencia social, tras seis meses, un año, y más. Impacto en sus vidas.
Resultado del programa	Efectividad de programas específicos en conseguir los resultados deseados	Tasa de limpieza de las calles, resultado de barrer	Número de gente colocada en trabajos, trabajando y fuera de la asistencia social, tras seis meses, un año, y más. Impacto en sus vidas.
Resultado político	Efectividad de políticas más amplias a la hora de conseguir metas.	Medidas que indiquen cuánta basura tiran los ciudadanos en la calle	Porcentaje de fuerza de trabajo potencial desempleada, en la asistencia social, y pobre; tanto por ciento de la población en la asistencia social, que está en ella más de un año, cinco años, etc.
Eficacia del programa	Coste del resultado por unidad	Coste por kilómetro de calle barrida	Coste por trabajo formado; colocación; trabajo conservado; etc.

Anticipar resistencias poderosas. La información completa sobre la eficacia y la efectividad puede ser muy amenazadora para los proveedores de servicios que dudan de su habilidad para competir. Florida revocó un programa que premiaba a las escuelas individuales para mejorar los rendimientos. Arizona dejó sin fondos un programa que hacía públicas las tasas de colocación de los graduados de todas las instituciones de formación y educación de postsecundaria. Los *colleges* y las escuelas técnicas de la comunidad encontraron la información muy amenazadora, puesto que revelaba su auténtica eficacia a la hora de preparar a la gente para los trabajos reales. «Tienes grandes ganadores y grandes perdedores», afirma George Britton, que colaboró en el diseño del sistema mientras servía a las órdenes del gobernador Bruce Babbitt. «Cuando, en una función gubernamental, singularizas a gente arriba y abajo, es algo muy amenazador.» (La publicación de estos datos no es, sin embargo, imposible. Florida lo hace actualmente con éxito.)

Hacer participar a proveedores y empleados a la hora de desarrollar las medidas adecuadas. El mejor modo de afrontar las resistencias es unir a los proveedores y los empleados en el proceso de definir las medidas adecuadas. Después de todo, para utilizar los datos con eficacia la gente ha de confiar en su valor. Necesita ser «dueña» de las medidas específicas que se utilicen, para tener la sensación de que proporciona información útil y relevante que mejorará el servicio que presta. Aquellos que se oponen, o se oponen a medidas concretas, precisan de una justa oportunidad. El hacer enfadar a la gente con medidas inadecuadas en cuyo desarrollo no han participado es un modo seguro de crear resistencias, destruir la moral, y estimular las trampas.

Someter las medidas a revisiones y modificaciones anuales. Ninguna medida es perfecta. Como gobernar no es una ciencia, es imposible aislar medidas que reflejen perfectamente los resultados de la actividad de gobierno. Todo lo que podemos hacer es esperar una aproximación, a menudo utilizando los poderes que mejor tengamos a mano. Por lo tanto tiene sentido modificar y ajustar con frecuencia las medidas de los rendimien-

Eficacia política	Coste de conseguir las metas fundamentales	Coste por nivel X de limpieza de la calle	Coste de conseguir la disminución deseada en el desempleo, la tasa de pobreza, carga del bienestar, etc.
Efectividad del programa	Grado en el que el programa rinde los resultados deseados	Nivel de satisfacción del ciudadano con la limpieza de sus calles	Número de personas colocadas, trabajando, y fuera de la asistencia social después de seis meses, un año, y más; impacto en sus vidas.
Efectividad política	Grado en el que concuerdan las necesidades de los ciudadanos y las metas fundamentales	¿Quieren los ciudadanos utilizar el dinero de este modo? Por ej., ¿lo prefieren gastar en pavimentar las calles?	Efecto en la sociedad en su conjunto: por ej., tasa de pobreza, carga en la asistencia social, tasa de criminalidad, próximos gastos para remediar la pobreza, etc.

Vigilar la discriminación. Los proveedores de servicios normalmente proporcionan las cantidades que se les piden, incluso si para hacerlo tienen que ahorrar. Si tienen que colocar 1.000 personas al año, buscarán a las 1.000 personas más empleables que puedan encontrar y las formarán; una práctica conocida como «hacer crema». Esto es precisamente lo que sucedió en los primeros años del Job Training Partnership Act (Decreto de la Asociación de Formación para el Trabajo). Una solución es establecer metas para los diferentes grupos de población: 250 plazas para aquellos que han estado sin trabajo por lo menos durante dos años; 250 para aquellos que no tienen estudios superiores; etc. Otra es establecer distintas tasas de reembolsos para poblaciones distintas. La industria de formación del trabajador ha aprendido a utilizar dichos mecanismos sin grandes dificultades.

tos, en particular si sus defectos han aparecido en la práctica. También tiene sentido hacer periódicamente una auditoría de las medidas independientes, para ver si éstas en realidad miden lo que uno cree que miden. (Una vez las medidas se han ajustado, sin embargo, es mejor dejar algunas medidas un tiempo, para que sirvan para comparar los rendimientos de un año para el otro.)

No utilizar demasiadas o demasiado pocas medidas. Si una organización establece demasiado pocas medidas, puede no reflejar todas sus metas. Por lo que sus proveedores de servicios pueden verse impulsados a acentuar algunas metas a expensas de otras. Si establece demasiadas, puede diluir el poder de todas las medidas. Los proveedores pueden confundirse sobre las prioridades y verse cargados de papeleo, y los gestores verse superados por los detalles. Si los empleados y los proveedores de servicios participan en el desarrollo de las medidas, y si se les permite corregirlas periódicamente, serán capaces de hallar el equilibrio correcto.

Huir de incentivos erróneos. La fórmula de financiación utilizada para financiar la asistencia en casa, durante los años 70, creó en Illinois, como hemos visto en el capítulo 5, un incentivo para mantener a la gente enferma en cama. Las medidas de actuación usadas originalmente por la JTPA alentaron a los proveedores de formación de trabajo a discriminar. Los incentivos erróneos como éste pueden minar el esfuerzo global de la medición de actuaciones. Para eliminarlos, las organizaciones deben «jugar» con nuevas medidas —anticipando cómo responderán los proveedores listos— antes de imponerlas.

Mantener la la función de control y medida en una oficina imparcial, políticamente independiente. Si la gente se ha de fiar de los datos, deben confiar en su objetividad. Por lo tanto es una buena idea utilizar una oficina independiente, como la Phoenix City Auditor's Office (Oficina de Auditores de la Ciudad de Phoenix), para llevar a cabo las medidas. En Florida, dos organizaciones privadas, la Florida TaxWatch (Control de Impuestos de Florida) y el Florida Council of 100 (Consejo de Florida de los 100) (un grupo empresarial) han creado una organización no gu-

bernamental, Partners in Productivity (Socios en Productividad), para desarrollar medidas de control de rendimiento para el gobierno del Estado. En Gran Bretaña la Comisión Nacional de Auditoría controla las actuaciones no sólo de las agencias del gobierno nacional, sino también de los gobiernos locales. Puesto que publica información comparativa sobre la eficacia y la efectividad, los gobiernos locales prestan mucha atención a sus estudios.

Concentrarse en maximizar el uso de los datos de rendimiento. El simple desarrollar medidas no garantiza que los gestores las utilicen para cambiar lo que hacen o que las legislaturas las utilicen para cambiar lo que financian. El Fondo para la Ciudad de Nueva York descubrió, según Greg Farrell, que «las buenas medidas y la información de la gestión son más fáciles de concebir que de integrar en la conducta de las empresas del gobierno... los gestores del gobierno, la mayoría, no *están acostumbrados* a tener o utilizar la información sobre gestión, en especial para propósitos de futura prospección. En muchos casos, las presiones políticas son a menudo tan grandes que los *datos* parecen estar fuera de lugar a la hora de tomar decisiones».⁴ Por lo tanto mientras desarrollan medidas de rendimiento, las organizaciones han de intentar desarrollar presupuestos, sistemas de gestión y sistemas de recompensa basados en los datos sobre rendimiento, como hemos discutido en el capítulo 5.

4. Gregory Farrell, «Helping Government Do Better with Less», en *Fund for the City of New York, Three-Year Report, 1979-81*, págs. 5-6.