

Evaluaciones que mejoran el desempeño de las intervenciones públicas en América Latina y el Caribe

Mauricio García Moreno

Mauricio García Moreno

Consultor independiente. Licenciado en Antropología y máster en Administración de Empresas. Tiene amplia experiencia en las áreas de investigación social, diseño y análisis de políticas públicas, desarrollo institucional, seguimiento y evaluación. Ha abordado distintas dimensiones de las intervenciones para el desarrollo: estudios de la problemática social e institucional, estudio y diseño de marcos normativos e institucionales para crear valor público, seguimiento y evaluación de programas y proyectos, capacitación y gestión del conocimiento. Desde hace diez años su trabajo se enfoca en el desarrollo de intervenciones e instrumentos que mejoren la capacidad de las instituciones nacionales y subnacionales para implementar una gestión orientada a resultados, con énfasis en las áreas de presupuesto basado en el desempeño, seguimiento y evaluación. Sus trabajos en este campo fueron publicados por el Banco Interamericano de Desarrollo, donde fue consultor senior. Actualmente se desempeña como consultor internacional independiente.

Las comunicaciones con el autor pueden dirigirse a:
E-mail: maurgarm@gmail.com

Evaluaciones que mejoran el desempeño de las intervenciones públicas en América Latina y el Caribe

Este artículo analiza las características de las evaluaciones rápidas, un tipo particular de evaluación de las intervenciones públicas que se ha difundido en la última década en varios países de América Latina. En la mayoría de casos, estas evaluaciones forman parte de las reformas impulsadas por los entes rectores del presupuesto nacional y están encaminadas a mejorar la eficacia y eficiencia del sector público. El artículo examina el proceso de gestión de las evaluaciones rápidas -que comprende la selección de las intervenciones que se examinarán, la ejecución de las evaluaciones, y el uso de los resultados- en Chile, Colombia, México, Paraguay, Perú y Uruguay. Finalmente, ofrece recomendaciones para mejorar la efectividad de estos instrumentos.

Palabras clave: Evaluación; Gestión Pública; Presupuesto; Rendimiento; Mejoramiento del Rendimiento; Gestión por Resultados; América Latina; Caribe; Chile; Colombia; México; Paraguay; Perú; Uruguay

Evaluations that Improve the Performance of Public Interventions in Latin America and the Caribbean

This paper analyzes the characteristics of rapid evaluations, a particular type of evaluation of public interventions that has become widespread in the last decade in several Latin American countries. In most cases, these evaluations are part of the reforms promoted by the national budget governing bodies and aimed at improving the effectiveness and efficiency of the public sector. The article examines the management process of rapid assessments -which includes the selection of interventions to be examined, the execution of the assessments, and the use of the results- in Chile, Colombia, Mexico, Paraguay, Peru, and Uruguay. Finally, it offers recommendations to improve the effectiveness of these tools.

Key words: Evaluation; Public Management; Budget; Performance; Performance Improvement; Management by Results; Latin America; Caribbean; Chile; Colombia; Mexico; Paraguay; Peru; Uruguay

Recibido: 26-07-2021 y 27-10-2021 (segunda versión) Aceptado: 27-10-2021.

Introducción

Fortalecer la capacidad para evaluar la acción de las intervenciones públicas¹ es parte de la agenda de todas las reformas conducentes a mejorar la administración gubernamental. Las iniciativas que procuran mejorar la gestión del desempeño, implementar un presupuesto por resultados o crear unidades que impulsen las prioridades fijadas por el gobierno, requieren de un modo u otro evaluar las acciones que las instituciones realizan para satisfacer las demandas de los ciudadanos o para cumplir las promesas de los gobernantes. De la misma manera, los propósitos de mejorar los servicios públicos o lograr más transparencia en la gestión del Estado son alcanzables solo en la medida en que sea posible evaluar el desempeño de las políticas, organizaciones y programas.

A pesar de su centralidad, la capacidad de evaluar las intervenciones públicas es una de las asignaturas pendientes de la mayoría de los países de América Latina y el Caribe (ALC). Solo pocos realizan acciones sistemáticas y perseveran en el afán de dotarse de capacidades que les permitan valorar la pertinencia, eficacia, eficiencia, calidad y sostenibilidad de las acciones que los gobiernos llevan a cabo para mejorar la vida de los ciudadanos (García Moreno, Kaufmann y Sanginés, 2015; CEPAL, 2018; Pérez Yarahuán y Maldonado, 2015).

La adopción generalizada durante este siglo de indicadores de desempeño para hacer seguimiento a la ejecución del presupuesto y/o a los planes de gobierno, por lo general no se ha complementado con iniciativas vigorosas que fomenten la evaluación de las intervenciones públicas, lo que limita seriamente la utilidad de esos indicadores, ya que por sí solos son insuficientes para dar cuenta de qué funciona en una intervención pública, cómo opera, qué efectos produce y por qué. Esto debido, entre otras razones, a que el desempeño es el resultado de múltiples variables que solamente un estudio evaluativo puede sopesar adecuadamente.

Algunos países de la región, inspirados en el ejemplo de la Dirección de Presupuestos de Chile (DIPRES) primero y del Consejo Nacional de Evaluación de la Política Social (CONEVAL) de México después, han dado pasos hacia la construcción de sistemas de evaluación que paulatinamente se han instalado como mecanismos útiles para mejorar la gestión pública. Colombia lleva más de una década aplicando sistemáticamente distintos tipos de evaluaciones a las intervenciones consignadas en el Plan Nacional de Desarrollo.

Perú y, más recientemente, Paraguay y Uruguay también están empeñados en realizar evaluaciones de sus intervenciones públicas con el propósito no solamente de rendir cuentas, sino también de mejorar su efectividad. En todos los casos, el fortalecimiento de la capacidad institucional para evaluar tuvo lugar al amparo de iniciativas orientadas a implementar la gestión para resultados o el presupuesto por resultados².

Aunque incipiente, esta corriente que busca fortalecer la evaluación es significativa por dos razones: 1) tiene un carácter marcadamente institucional y sistemático, porque incluye la evaluación en el ciclo de las políticas públicas o del presupuesto y asigna a una entidad de alta jerarquía la responsabilidad de gestionarla; y 2) los estudios de evaluación no son considerados como un medio de rendición de cuentas solamente, sino también como mecanismo para mejorar la gestión de las intervenciones públicas y la toma de decisiones.

Los países que han incorporado la evaluación de esta manera, también adoptaron las evaluaciones rápidas de las intervenciones públicas (ERIP) como su principal instrumento de estudio. Este tipo de evaluación tiene la virtud de resolver dos inconvenientes que las evaluaciones de otro tipo -como las de impacto, las de costo-eficiencia o costo-beneficio- presentan: el costo y el tiempo de ejecución. En efecto, las ERIP se realizan en lapsos cortos (entre cuatro y ocho meses), a costos bajos (menos de 80.000 dólares), y usando la información que ya tienen los programas: indicadores de desempeño, información contable, registros administrativos, etc. Además, se ajustan al ciclo anual del presupuesto y, por lo tanto, la información que producen puede usarse durante las etapas de formulación y discusión.

El objetivo de este estudio es analizar las características de las ERIP que se realizan en ALC -como parte de los esfuerzos sistemáticos que los gobiernos llevan a cabo para mejorar el desempeño de sus intervenciones- con el propósito de identificar los elementos que deben tomarse en cuenta para aplicarlas o para optimizar su funcionamiento en aquellos países que ya las emplean. Los estudios publicados sobre la materia y los documentos oficiales existentes (normas, metodologías, informes) se usaron como fuentes de información. El texto está organizado en cuatro secciones. La primera es una introducción al tema y destaca la importancia de la evaluación para la gestión de las políticas públicas. La segunda presenta las características de las

Las evaluaciones rápidas surgieron como respuesta a la necesidad de los gestores y supervisores de los programas y políticas públicas de abreviar el proceso de investigación y reducir el costo de las evaluaciones tradicionales.

ERIP que se llevan a cabo en seis países de ALC que actualmente realizan este tipo de evaluación de manera sistemática y centralizada: Chile, Colombia, México, Paraguay, Perú y Uruguay. La tercera sección analiza el proceso de gestión de las ERIP, que empieza con la selección de las intervenciones públicas, pasa por la preparación y ejecución de los estudios de evaluación, y termina con el uso de los resultados que arrojan las evaluaciones. La cuarta sección brinda sugerencias a los países que realizan ERIP para mejorar su gestión y uso; también ofrece recomendaciones a los países que aún no las llevan a cabo y que están interesados en hacerlo.

1. Las ERIP y su aplicación en América Latina y el Caribe

Las evaluaciones rápidas surgieron como respuesta a la necesidad de los gestores y supervisores de los programas y políticas públicas de abreviar el proceso de investigación y reducir el costo de las evaluaciones tradicionales, que se enfocaban en identificar los efectos de las intervenciones públicas mediante procedimientos experimentales (Anker ...[et al], 1993). Las evaluaciones que se realizaban hasta los años setenta en los Estados Unidos, por ejemplo, demandaban varios años de trabajo y consumían una gran cantidad de recursos (Weiss, 1998). En muchos casos, esto dificultaba el uso de los hallazgos para planificar, asignar recursos y corregir la gestión de los programas evaluados, porque los resultados de los estudios no estaban disponibles oportunamente.

A partir de la década de los ochenta, varias organizaciones empezaron a ensayar métodos de evaluación más rápidos. El Alto Comisionado de las Naciones Unidas para los Refugiados (ACNUR) desarrolló un método para evaluar la efectividad de sus acciones de respuesta a las emergencias humanitarias (*Real Time Evaluations*). La Organización Mundial de la Salud creó un método para examinar la calidad de los servicios de salud, identificar problemas operativos y proveer información a los administradores para que tomen medidas correctivas (*Rapid Evaluation Method*). Varias organizaciones y expertos vinculados a programas de desarrollo rural crearon la metodología *Evaluación Rural Participativa*, que sigue siendo un método muy difundido para evaluar intervenciones en agricultura, medio ambiente y reducción de la pobreza (McNall y Foster-Fishman, 2007).

¿Qué distingue a las evaluaciones rápidas de otras evaluaciones?

Hargreaves (2014) propone que aunque los métodos de evaluación rápida son heterogéneos en cuanto a su objetivo y complejidad, existen tres principios que los homogeneizan: 1) buscan un equilibrio entre el análisis de los productos y de los resultados; 2) no son instrumentos de medición o diagnóstico únicamente, sino parte del proceso de gestión; y 3) promueven que la información que se recolecta y analiza sea usada como un catalizador del cambio continuo.

Por otro lado, McNall y Foster-Fishman (2007) sostienen que las evaluaciones rápidas comparten las siguientes características. Desde el punto de vista del método, usan diversas técnicas de recolección de información, aplican métodos cuantitativos y cualitativos, y usan datos que los programas ya tienen. Desde el punto de vista del proceso, generalmente incluyen la participación de la población, involucran de distintas formas a los miembros del equipo del programa evaluado, y se realizan de manera continua en un ciclo conformado por la recolección, el análisis y el uso de la información.

El desarrollo de la informática y, específicamente, la creación de sistemas de información para la gestión (*Management Information Systems*) y de sistemas integrados de información financiera han contribuido a fortalecer y ampliar el uso de las evaluaciones rápidas, en la medida en que proporcionan registros administrativos abundantes sobre los insumos, procesos y productos de los programas y organizaciones. Además, los sistemas informáticos actuales permiten procesar y analizar la información con gran rapidez y de forma descentralizada. Anteriormente, la clasificación y/o el procesamiento de esos datos consumía mucho tiempo y formaba parte del proceso evaluativo.

La institucionalización de las ERIP para el estudio sistemático de las intervenciones públicas tuvo uno de sus hitos mundiales en la implementación del *Program Assessment Rating Tool* (PART) entre 2003 y 2008 por parte del gobierno de los Estados Unidos. La entidad a cargo de esta tarea fue la *Office of Management and Budget* (OMB), dependiente de la presidencia, y usó el PART para examinar más de mil programas que emplearon el 98% de los recursos del presupuesto del gobierno federal en un lapso de 5 años (Moynihan, 2013). El instrumento contenía entre 25 y 30 preguntas divididas en cuatro categorías: propósito y diseño del programa, planificación estratégica, gestión del programa y resultados del programa. Sobre la base de las respuestas, la OMB

Las ERIP que se llevan a cabo en ALC tienen dos propósitos bien definidos: 1) mejorar el diseño, calidad, eficiencia y eficacia de las intervenciones públicas; y 2) informar a los actores que intervienen en el ciclo presupuestario acerca del desempeño de las intervenciones públicas.

clasificaba a los programas en efectivos, moderadamente efectivos, adecuados o inefectivos. El PART fue un referente importante para ALC, las ERIP desarrolladas en Colombia y México se inspiraron en parte en el modelo creado por la OMB (CONEVAL, 2015; DNP, 2010).

Características de las ERIP que se realizan en ALC

Las ERIP que se llevan a cabo en ALC tienen dos propósitos bien definidos: 1) mejorar el diseño, calidad, eficiencia y eficacia de las intervenciones públicas; y 2) informar a los actores que intervienen en el ciclo presupuestario acerca del desempeño de las intervenciones públicas. El segundo objetivo está presente solamente en aquellos países cuyo sistema de evaluación es parte de la implementación del presupuesto por resultados.

Los productos que generan las ERIP son tres: 1) el informe de los estudios de evaluación, 2) el diálogo técnico sobre el desempeño que mantienen los actores involucrados en las intervenciones públicas que se evalúan, y 3) el plan de acción que la entidad ejecutora de la intervención pública acuerda con el ente rector de las evaluaciones como consecuencia del análisis de los hallazgos y recomendaciones de los evaluadores.

Para lograr los objetivos señalados, los organismos encargados de las evaluaciones han establecido normas, procedimientos e instrumentos que son propias de las ERIP y que las diferencian de otros tipos de evaluación. A continuación, se analizan las características de las ERIP que se enuncian de manera resumida en la Tabla 1.

Tabla 1
Principales características de las ERIP

Ámbito	Características
Marco legal	Normas que establecen con detalle el rol y las características de los sistemas de seguimiento y evaluación en la implementación de la gestión para resultados y/o el presupuesto basado en el desempeño.
Gestión	Centralizada en el ente rector de planificación o presupuesto.
Objeto de análisis	Programas, subprogramas e instituciones.
Programación	Plan anual.
Metodología	Estandarizada en guías, manuales o términos de referencia.

Tabla 1 (continuación)
Principales características de las ERIP

Ámbito	Características
Duración	Entre 4 y 8 meses.
Análisis de la cadena de valor	Diseño, implementación y resultados inmediatos.
Fuentes de información	Indicadores de desempeño, registros administrativos, entrevistas.
Ejecutor de la evaluación	Evaluador independiente.

Fuente: elaboración propia.

Marco legal: las ERIP se ejecutan para cumplir disposiciones legales que regulan detalladamente el proceso de evaluación. Estas disposiciones forman parte de normas más amplias que estipulan que la gestión pública debe estar orientada al logro de resultados. El marco legal establece una institución rectora o responsable del sistema de evaluación.

Gestión de las evaluaciones: la gestión de las ERIP es centralizada. Esto significa que la entidad que toma las decisiones clave de la evaluación -qué, cómo y cuándo evaluar- es distinta a la que ejecuta la intervención pública. Por lo general, esa entidad es el ente rector del presupuesto o de la planificación que, además, se encarga de la contratación de las evaluaciones.

Objeto de análisis: las ERIP analizan proyectos, subproyectos, programas o componentes de programas, que aquí se denominan con el término genérico de intervenciones públicas. También analizan el desempeño de toda una organización. Las ERIP no evalúan políticas públicas.

Programación: las ERIP se programan mediante un plan anual en el que se coordinan el proceso de ejecución de las evaluaciones (diseño, ejecución, informe, compromisos de mejora) con el ciclo presupuestario (programación, formulación, aprobación, ejecución, control y evaluación). Se considera que la coordinación entre esos dos procesos es necesaria para asegurar el uso de los informes de evaluación en el ciclo presupuestario.

Metodología: las ERIP usan metodologías estandarizadas en guías, manuales o términos de referencia con el propósito de homogeneizar los procesos de evaluación. El grado de detalle de las metodologías varía de país a país y va desde un listado de variables

que los evaluadores deben considerar para su ejercicio hasta un instrumento estructurado mediante preguntas cerradas.

Duración: la duración de las evaluaciones varía entre cuatro y ocho meses. Este es uno de los elementos clave de este tipo de evaluación que condiciona otras características como las fuentes de información y los criterios de evaluación que se aplican.

Análisis de la cadena de valor: las ERIP analizan fundamentalmente el diseño y la implementación (insumos, procesos y productos) de las intervenciones públicas, aunque algunas incluyen el análisis de los resultados inmediatos (efectos del uso de los bienes y servicios en los ciudadanos). Las ERIP no tienen el objetivo de evaluar el impacto de las intervenciones públicas ni de pronunciarse sobre las causas y atribución de los resultados medios y finales.

Fuentes de información: la principal fuente de información de las ERIP son los indicadores de desempeño y los registros administrativos de las intervenciones públicas que se evalúan. Algunas evaluaciones complementan esta información con entrevistas a informantes calificados y/o con encuestas.

Ejecutor de la evaluación: por lo general los ejecutores de los estudios de evaluación son evaluadores independientes contratados mediante concurso público, aunque en algunos casos también participan funcionarios de la entidad rectora de las evaluaciones o de la que ejecuta la intervención.

Países de ALC que ejecutan ERIP

Los países de ALC que realizan sistemáticamente ERIP son: Chile, Colombia, México, Paraguay, Perú y Uruguay. Para ello han creado sistemas de evaluación compuestos por normas, metodologías e instrumentos que funcionan con recursos financieros y humanos en un marco institucional ligado a las funciones de planificación y/o presupuesto nacional (Tabla 2). La mayoría de estos países realizan diversos tipos de evaluaciones, generalmente ERIP y evaluaciones de impacto.

En la lista no constan Brasil ni Costa Rica, países que según varios estudios tienen capacidades institucionales de evaluación superiores al promedio de ALC (García Moreno y Feinstein, 2015; Pérez Yarahuan y Maldonado, 2015). En Brasil, los ministerios sectoriales -especialmente el Ministerio de Desarrollo Social- realizan evaluaciones de distinto tipo a los programas sociales, pero estas no cumplen con los criterios sobre el marco legal, la gestión y la programación

enunciados anteriormente. Lo mismo ocurre en Costa Rica, aunque el Ministerio de Planificación Nacional y Política Económica (MIDEPLAN) tiene experiencia en la gestión de evaluaciones, su sistema de evaluación se basa en el diseño descentralizado y *ad hoc* de términos de referencia y no en formatos estandarizados que se gestionan centralmente (MIDEPLAN, 2012 y 2017).

Tabla 2
Ente rector y marco legal de las ERIP

País	Ente Rector	Marco legal ³
Chile	Presupuesto (Ministerio de Hacienda)	2003. Decreto 1.177. Reglamento para el funcionamiento del sistema de evaluación de programas establecido en el Artículo 52 del Decreto Ley 1.263 de 1975, Orgánico de Administración Financiera del Estado.
Colombia	Planificación (Departamento Nacional de Planeación)	2014. Decreto 1.290 por el cual se reglamenta el Sistema Nacional de Evaluación de Gestión y Resultados (SINERGIA).
México	Órgano independiente (Consejo Nacional de Evaluación de la Política Social)	2007. Lineamientos generales para la evaluación de los programas federales de la Administración Pública Federal.
Paraguay	Presupuesto (Ministerio de Hacienda)	2017. Decreto 6.715/2017, Guía de Normas y Procesos de Ejecución de la Ley 5.554/2016 Vigente para el Ejercicio Fiscal 2017. Anexo A.
Perú	Presupuesto (Ministerio de Economía y Finanzas)	2008. Directiva 009-2008-EF/76.01 para las evaluaciones independientes en el Sistema Nacional de Presupuesto Público en el marco del Presupuesto por Resultados (modificada por la Resolución Directoral 023-2012-EF/50.01).
Uruguay	Planificación (Oficina de Planeamiento y Presupuesto de la República)	2011. Ley 16.736 modificada por la Ley 18.996 de 2011.

Fuente: elaboración propia.

Clasificación de las ERIP

Se ha clasificado a las ERIP que se realizan en ALC de acuerdo a tres criterios: objeto de la evaluación, unidad de análisis y tipo de instrumento (Tabla 3).

Objeto de la evaluación

El objeto de la evaluación es el elemento de la cadena de valor cuyo desempeño se examina: diseño, procesos, productos y resultados. De acuerdo con este criterio las evaluaciones se clasifican en dos tipos: aquellas que incluyen los cuatro elementos y aquellas que incluyen solamente algunos de ellos. Un ejemplo de las primeras son las *Evaluaciones de Programas Gubernamentales* de Chile y las *Evaluaciones de Consistencia y Resultados* de México. Un ejemplo de las segundas son las *Evaluaciones de Procesos de México*. De los 17 instrumentos de evaluación analizados, 6 corresponden al primer tipo y 11 al segundo.

La mayoría de países empezaron sus sistemas de evaluación con evaluaciones del primer tipo y, posteriormente, algunos han ido creando instrumentos con un enfoque menos abarcador y más desagregado, como se observa en Colombia y Paraguay. El Ministerio de Economía y Finanzas (MEF) de Perú anunció en 2018 que dividiría las *Evaluaciones de Desempeño y Ejecución Presupuestal* (EDEP) en evaluaciones de diseño y de procesos⁴. En Paraguay, las razones de este cambio fueron tres: 1) enfocar los ejercicios evaluativos en los aspectos que el Ministerio de Hacienda considera más relevantes en cada uno de los programas presupuestarios; 2) no desperdiciar esfuerzos en evaluar resultados en programas que no tienen información adecuada; y 3) reforzar la evaluación del diseño de los programas en miras a consolidar la base programática del presupuesto por resultados. Por otro lado, un estudio comisionado por el MEF de Perú sobre las EDEP (Baanante y Gracia, 2015) encontró que una de sus debilidades era la inflexibilidad del instrumento, poco adaptable a la variedad de características y circunstancias de las intervenciones públicas. El estudio sugirió mantener la estandarización del formato, pero introducir elementos que permitan que las evaluaciones se ajusten mejor a su objeto de evaluación.

Chile, por otro lado, se ha mantenido fiel al enfoque abarcador de sus evaluaciones. La evaluación realizada en 2005 (Banco Mundial, 2005) a las *Evaluaciones de Programas Gubernamentales* (EPG) y otros análisis posteriores (DIPRES, 2016) encontraron que

La ventaja de las evaluaciones enfocadas en alguno de los elementos de la cadena de valor (Colombia, México y Paraguay) es que permiten a los entes que evalúan adaptar con más facilidad sus instrumentos a las características y necesidades de las intervenciones públicas.

el enfoque amplio de las EPG -basado en el marco lógico- es útil para los propósitos del ente evaluador. Incluso la última metodología de evaluación creada por la DIPRES, que se aplica desde 2014 (Evaluación Focalizada de Ámbito -EFA), incluye varias dimensiones: diseño, implementación y costos.

La ventaja de las evaluaciones enfocadas en alguno de los elementos de la cadena de valor (Colombia, México y Paraguay) es que permiten a los entes que evalúan adaptar con más facilidad sus instrumentos a las características y necesidades de las intervenciones públicas. Así, por ejemplo, es posible que el Ministerio de Hacienda esté interesado en evaluar el diseño de un programa, pero no sus resultados, porque no existen datos confiables para analizarlos ni tampoco recursos para recopilarlos.

Unidad de análisis

Respecto a la unidad de análisis, las evaluaciones se clasifican en dos tipos: las que analizan programas e intervenciones públicas y las que analizan organizaciones. El primer tipo es el más difundido, 15 de los 17 instrumentos tienen ese enfoque. Solamente Chile y Colombia realizan evaluaciones a organizaciones. En ambos países se examina la capacidad institucional para usar los recursos a su disposición, coordinar su acción con otras entidades, y cumplir los objetivos trazados por las políticas públicas. En los países que aplican una estructura programática al presupuesto, la unidad de análisis de las evaluaciones es el programa presupuestario. La evaluación de organizaciones requiere un grado de madurez mayor respecto a la capacidad institucional y autoridad de los entes evaluadores, los sistemas de información de los registros administrativos, la información estadística, la gestión de las políticas públicas, entre otros aspectos.

Tipos de instrumentos de evaluación

Respecto al tipo de instrumento que usan las ERIP, se sigue la clasificación sugerida por Feinstein (2016): 1) evaluaciones basadas en un cuestionario con preguntas sobre las áreas del desempeño; 2) evaluaciones guiadas por un marco general sobre las áreas de desempeño; y 3) evaluaciones que combinan los dos modelos anteriores.

El tipo de instrumento más extendido es el cuestionario (8 de los 17 instrumentos lo usan), seguido por el marco general (5) y el modelo mixto (4). La ventaja del cuestionario sobre los otros tipos es que disminuye la complejidad de la gestión de las evaluaciones

en actividades tales como el diseño de los términos de referencia y la revisión de los informes. Además, en el caso de las evaluaciones de México y Paraguay, permite elaborar una matriz cuantitativa que resume la información cualitativa.

Las EPG de Chile son un caso interesante de modelo mixto, porque parten de un marco general que se complementa con pautas detalladas para la elaboración del informe de evaluación (DIPRES, s.f.) y con documentos técnicos que precisan los conceptos e instrumentos que deben utilizarse (DIPRES, 2015 y 2007).

En general, se observa que los instrumentos de todos los países proporcionan a los evaluadores instrucciones precisas de lo que deben evaluar y la forma de hacerlo, dejando poco margen al arbitrio, no solo de los evaluadores, sino también de quienes elaboran los términos de referencia.

Tabla 3
Características de los instrumentos de evaluación

País	Nombre del instrumento	Año de inicio	Objeto de evaluación	Unidad de análisis	Tipo de instrumento
Chile	Evaluaciones de Programas Gubernamentales	1997	DI / PC / PD / RE	PRG	MX
	Evaluación del Gasto Institucional	2002	DI / PD / RE	INS	GE
	Evaluación Focalizada de Ámbito	2014	DI / PC / PD	PRG	MX
Colombia	Evaluación Ejecutiva ⁵	2006	DI / PC / PD / RE	PRG	CU
	Evaluación de Operaciones	2012	DI / PC / PD	PRG	GE
	Evaluación de Resultados	2012	PD / RE	PRG	GE
	Evaluación Institucional	2012	DI / PC	INS	GE
México	Evaluación de Diseño	2007	DI	PRG	CU
	Evaluación de Consistencia y Resultados	2007	DI / PC / PD / RE	PRG	CU
	Evaluación Específica de Desempeño	2009	PD / RE	PGR	CU
	Evaluación de Procesos	2009	PC	PRG	CU

Tabla 3 (continuación)
Características de los instrumentos de evaluación

Paraguay	Evaluación de Diseño y Gestión ⁶	2011	DI / PC / PD / RE	PRG	MX
	Evaluación de Eficiencia ⁷	2011	PC / PD	PRG	GE
	Evaluación de Diseño	2018	DI	PRG	CU
	Evaluación de Implementación	2018	PC / PD	PRG	CU
	Evaluación de Desempeño	2018	RE	PRG	CU
Perú	Evaluaciones de Desempeño y Ejecución Presupuestal	2008	DI / PC / PD / RE	PRG	MX
Uruguay	Evaluaciones de Diseño, Implementación y Desempeño	2011	DI / PC / PD / RE	PRG	CU
Leyenda	Objeto de evaluación: DI = diseño, PC = proceso, PD = producto, RE = resultado				
	Unidad de análisis: PRG = programa o intervención pública, INS = institución				
	Modelo: CU = cuestionario, MX = mixto, GE = general				

Fuente: elaboración propia.

Existen limitantes para valorar adecuadamente la cobertura de las ERIP, puesto que la información que proporcionan las entidades rectoras es incompleta e insuficiente.

Cobertura de las ERIP

Existen limitantes para valorar adecuadamente la cobertura de las ERIP, puesto que la información que proporcionan las entidades rectoras es incompleta e insuficiente. Por lo general, estas solamente ofrecen repositorios con los informes de las evaluaciones y cuadros estadísticos con el número de estudios realizados anualmente. Solo la DIPRES ha producido recientemente una metodología para medir la cobertura, es decir, “las evaluaciones de programas e instituciones que se han realizado en un determinado periodo de tiempo en relación con el total de programas e instituciones evaluables en el mismo periodo” (Darville, Díaz y Leiva, 2017: 16). De todas maneras, los datos que se presentan a continuación sobre el volumen de evaluaciones realizadas en cada país, proporcionan una idea acerca de la importancia de las ERIP en el conjunto de evaluaciones. Claramente, las ERIP son el tipo de evaluación que se realiza con más frecuencia.

Chile

La DIPRES realiza cinco tipos de evaluaciones: Evaluaciones de Programas Gubernamentales (EPG), Evaluación del Gasto Institucional (EGI), Evaluación Focalizada de Ámbito (EFA), Evaluación de Programas Nuevos (EPN) y Evaluaciones de Impacto (EI). Las tres primeras son de tipo ERIP. Entre los años 1997 y 2017 la DIPRES evaluó 583 programas. El 76,5% de ellos fueron ERIP, el 21,3% EI y el 2,2% EPN. La EPG es el tipo de estudio que más se realiza: 65,9% del total. La DIPRES ha evaluado anualmente un promedio de 4,9% del presupuesto en el periodo 1999-2017. Los programas u organizaciones que alguna vez han sido parte de un proceso evaluativo corresponden al 56% del presupuesto evaluable de 2017 (Darville, Díaz y Leiva, 2017)⁸.

Colombia

El primer tipo de ERIP que ejecutó Colombia fue la llamada *Evaluación Ejecutiva* que se implementó entre 2007 y 2011. Actualmente el Sistema Nacional de Evaluación de Gestión y Resultados (SINERGIA) ejecuta cuatro tipos de evaluaciones: Evaluación de Operaciones, Evaluación de Resultados, Evaluación Institucional -que son ERIP- y Evaluación de Impacto. De las 255 evaluaciones efectuadas entre 2001 y 2017, el 61% fueron ERIP y más de la mitad de ellas (56%) evaluaciones de resultados⁹.

México

México es el único país de ALC con un ente autónomo como órgano rector de las evaluaciones de los programas que ejecutan las entidades del sector social y no la entidad a cargo de la planificación o del presupuesto¹⁰. Entre 2007 y 2015 el CONEVAL realizó 1.763 evaluaciones de 10 tipos distintos. El 73,7% fueron ERIP e incluyen: Evaluación de Consistencia y Resultados, Evaluaciones Específicas de Desempeño, Evaluaciones de Procesos y Evaluaciones de Diseño. La mayoría de ERIP son Evaluaciones Específicas de Desempeño (54%) y Evaluaciones de Consistencia y Resultados (28,5%) (CONEVAL, 2015). Además, entre 2007 y 2019 la Secretaría de Hacienda y Crédito Público (SHCP) administró 411 evaluaciones de las cuales el 95% fueron ERIP¹¹.

Paraguay

Entre 2011 y 2017 el Ministerio de Hacienda realizó 34 evaluaciones de dos tipos: 28 Evaluaciones de Diseño y Gestión y 6 Evaluaciones

de Eficiencia. En 2018 reorganizó el sistema de evaluación y actualmente realiza tres tipos de ERIP que son complementarias entre sí: Evaluación de Diseño, Evaluación de Implementación y Evaluación de Desempeño.

Perú

En 2007 Perú inició la ejecución de Evaluaciones de Desempeño y Ejecución Presupuestal (EDEP), que corresponden al tipo ERIP, y Evaluaciones de Impacto de Presupuesto (EIP) (MEF, 2008). Entre 2008 y 2017 el Ministerio de Economía y Finanzas llevó a cabo 71 evaluaciones, 61 (85%) son EDEP y 10 son EIP (15%)¹². Entre 2015 y 2018 el porcentaje de presupuesto público evaluado con las EDEP y las EIP fue de 26,54% (MEF, 2018).

Uruguay

Entre 2012 y 2017 la Dirección de Gestión y Evaluación del Estado (AGEV) ejecutó 27 Evaluaciones de Diseño, Implementación y Desempeño (DID), que corresponde al tipo ERIP, y una evaluación de impacto¹³.

2. Ciclo de gestión de las ERIP

Aunque el ciclo de ejecución de las ERIP varía entre los países, en todos los casos se observan tres fases: 1) selección de las intervenciones públicas que se evaluarán, 2) preparación y ejecución de los estudios de evaluación, y 3) uso y difusión de los resultados de las evaluaciones. Además, en general el ciclo de ejecución de las ERIP se ajusta al ciclo presupuestario cuyas etapas son: programación, formulación, discusión y aprobación, ejecución y evaluación. Este vínculo es una de las características más importantes de las ERIP que se aplican como uno más de los instrumentos del Presupuesto por Resultados. A continuación, se analiza cada una de las etapas del ciclo de ejecución de las ERIP.

Selección de las intervenciones que se evaluarán

El primer paso en el proceso de ejecución de las ERIP es la selección de las intervenciones públicas que se evaluarán. Este paso generalmente se lleva a cabo el año fiscal anterior a aquel en el que se realizarán los estudios evaluativos y se materializa en un plan o una agenda de evaluaciones que se hace pública. En algunos países la agenda se publica con la ley anual de presupuesto.

En la mayoría de los países la selección de los programas la llevan a cabo el Poder Ejecutivo, la entidad que está a cargo del sistema nacional de evaluación -Ministerio o Secretaría de Hacienda y/o Planificación- y se cuenta con la participación de las entidades que ejecutan los programas. El grado en que las últimas intervienen es variable: en Colombia y Perú no se consulta a los ejecutores, mientras que en México la selección de las intervenciones públicas la realizan coordinadamente el CONEVAL y la SHCP. Las dos entidades cumplen distintas funciones en el sistema nacional de seguimiento y evaluación y ambas administran la ejecución de evaluaciones.

En Chile la selección de los programas se realiza mediante un proceso más complejo (Ministerio de Hacienda, 2003). En primer lugar un Comité Interministerial, conformado por el Ministerio Secretaría General de la Presidencia, el Ministerio de Desarrollo Social y el Ministerio de Hacienda (quien lo preside), analiza las necesidades de evaluación y elabora una lista que el Ministerio de Hacienda acuerda con la Comisión Especial Mixta de Presupuesto del Congreso Nacional. La participación del Congreso en esta etapa del proceso evaluador forma parte de los mecanismos de negociación del presupuesto anual entre el Poder Legislativo y Ejecutivo, que se conocen con el nombre de *Protocolos de Acuerdo*. Este mecanismo en muchos casos auxilia de manera efectiva la falta de reglas más formales que regulen la relación entre los dos poderes, no solo en cuanto a la elección de los programas a evaluar, sino también a otros aspectos presupuestarios (Arana Araya, 2013). Por otro lado, algunos estudios señalan que en la práctica la DIPRES elige los programas a evaluar de manera vertical, con escasa participación de los organismos ejecutores de las intervenciones públicas (Arancibia, 2018; Olavarría, 2012).

Tabla 4
Actores que intervienen en la selección de los programas que se evalúan

País	Actores
Chile	El Ministerio de Hacienda negocia con el Congreso Nacional la lista de programas a evaluar con base en las demandas que recibe de un Comité Interministerial
Colombia	Departamento Nacional de Planeación
México	Secretaría de Hacienda y Crédito Público y Consejo Nacional de Evaluación de la Política Social
Paraguay	Ministerio de Hacienda
Perú	Ministerio de Economía y Finanzas
Uruguay	Oficina de Planeamiento y Presupuesto

Fuente: elaboración propia.

En la mayoría de los países la lista de programas a evaluar (sean ERIP o de otro tipo) se oficializa en un documento que se denomina plan o agenda anual de evaluaciones. En México este documento contiene, además de la lista, indicaciones sobre el tipo de evaluaciones que se realizarán, los procedimientos que deben usarse, fechas de referencia para las etapas del ciclo de evaluación, entre otros aspectos previamente establecidos en las normas y leyes. En Chile la lista forma parte del Protocolo de Acuerdo entre el Ministerio de Hacienda y el Congreso Nacional sobre la Ley anual de Presupuesto o queda establecida en una resolución del Ministerio de Hacienda.

Preparación y ejecución de las evaluaciones

Una vez que se han seleccionado las intervenciones públicas que se evaluarán, el siguiente paso es la ejecución de las evaluaciones que generalmente involucra las siguientes actividades: elaboración de los términos de referencia, selección y contratación de los evaluadores, supervisión y realización del estudio de evaluación, y revisión de los informes. La ejecución de las evaluaciones está a cargo de las entidades que administran el sistema nacional de evaluación de cada país, según consta en la Tabla 5.

Tabla 5
Entidades que administran los sistemas nacionales de evaluación por país

País	Administradores del sistema nacional de evaluación
Chile	División de Control de la Gestión Pública de la Dirección de Presupuestos del Ministerio de Hacienda.
Colombia	Dirección de Seguimiento y Evaluación de Políticas Públicas del Departamento Nacional de Planeación (DNP).
México	La Unidad de Evaluación del Desempeño de la Secretaría de Hacienda y Crédito Público (SHCP) y la Dirección General Adjunta de Evaluación del Consejo Nacional de Evaluación de la Política Social (CONEVAL).
Paraguay	Coordinación de Monitoreo y Evaluación del Gasto Público de la Dirección General de Presupuesto del Ministerio de Hacienda.
Perú	Dirección de Calidad del Gasto Público de la Dirección General de Presupuesto Público del Ministerio de Economía y Finanzas (MEF).
Uruguay	Dirección de Gestión y Evaluación (AGEV) de la Oficina de Planeamiento y Presupuesto.

Fuente: elaboración propia.

Aunque las normas establecen que la información sobre el desempeño se usará para promover la eficiencia en la asignación del presupuesto y mejorar la gestión de las intervenciones públicas, en la práctica los países estudiados se han enfocado en este último punto.

En la mayoría de los casos las entidades que administran las evaluaciones contratan a evaluadores independientes. En México los entes rectores de la evaluación pueden delegar la contratación a las entidades ejecutoras de las intervenciones. El proceso de contratación se realiza de acuerdo a las normas vigentes de contratación pública. En la etapa de ejecución, la participación de las entidades ejecutoras de las intervenciones públicas que se evalúan se limita a proporcionar la información necesaria para el estudio y facilitar las actividades logísticas que se requieran.

El uso de los resultados de las evaluaciones

La literatura acerca del presupuesto por resultados coincide en señalar que la información sobre el desempeño de las intervenciones públicas sirve para dos propósitos ligados a la calidad del gasto público (Allen y Tomassi, 2001): 1) mejorar la eficiencia en la asignación de recursos, es decir, la capacidad para distribuir los recursos con base en las prioridades del gobierno y la efectividad de los programas; y 2) promover la eficiencia técnica u operativa en el uso de recursos, es decir, la capacidad para implementar programas y prestar servicios al costo más bajo posible con el máximo beneficio. El marco legal que sustenta a las ERIP en los países estudiados también enuncia estos dos propósitos (CONPES y DNP, 2004; MEF, 2008; Ministerio de Hacienda, 2003; CONEVAL, SHCP Y SFP, 2007). Otros usos de la evaluación ligados a la gestión del desempeño, como el aprendizaje institucional, el control de la gestión y la rendición de cuentas (Van Dooren, Bouckaert y Halligan, 2010), aparecen con menos énfasis y frecuencia en las normas nacionales.

Aunque las normas establecen que la información sobre el desempeño se usará para promover la eficiencia en la asignación del presupuesto y mejorar la gestión de las intervenciones públicas, en la práctica los países estudiados se han enfocado en este último punto. Para eso han desarrollado procesos e instrumentos que promueven el uso de las evaluaciones en los procesos de mejora de la gestión de las intervenciones (eficiencia técnica). Solamente Chile ha desarrollado procedimientos para que los actores que participan en el ciclo presupuestario examinen la información sobre el desempeño analizada en las evaluaciones. Además, en ningún país se usa la información sobre el desempeño como elemento determinante en la asignación de los recursos. Esta tendencia se observa no solamente en los países estudiados de América Latina, sino también en

otras regiones e incluso entre los países desarrollados (Moynihan y Beazley, 2016; Schick, 2014)¹⁴.

El proceso para usar las evaluaciones con el propósito de mejorar la eficiencia y eficacia de las intervenciones públicas incluye los siguientes pasos: 1) los evaluadores elaboran recomendaciones basadas en los hallazgos de los estudios; 2) tomando en cuenta las recomendaciones de los evaluadores, el ente rector de las evaluaciones y el organismo ejecutor de la intervención pública acuerdan los cambios que se introducirán para mejorarla y un cronograma para llevarlos a cabo; 3) el ente rector de las evaluaciones y el organismo ejecutor hacen seguimiento a la implementación de los acuerdos; 4) el ente rector elabora un informe sobre el cumplimiento de los acuerdos, sea anualmente y/o al finalizar el plazo establecido en el cronograma.

Los acuerdos sobre los cambios que requieren los programas para mejorar su eficiencia y eficacia son el elemento central del uso de las evaluaciones y, en muchos casos, la justificación misma del ejercicio evaluador. Todos los países estudiados implementan un esquema similar al descrito, cuya aplicación es posible gracias a que la gestión de las evaluaciones está centralizada en los entes rectores del presupuesto o la planificación, entidades que tienen la autoridad necesaria para vigilar que se cumplan los acuerdos.

No obstante, el grado de detalle de las normas e instrumentos que regulan este proceso, la rigurosidad con que se llevan a cabo las actividades, y el grado de autoridad de los entes que gestionan la evaluación, son elementos que influyen en la aplicación efectiva de los cambios establecidos en los acuerdos. El CONEVAL de México y la DIPRES de Chile exhiben los procesos más sólidos de definición, adopción y seguimiento de los acuerdos mediante los instrumentos llamados Aspectos Susceptibles de Mejora (CONEVAL, 2011) en el primer país y los Compromisos Institucionales (Ministerio de Hacienda, 2003) en el segundo.

Otro uso importante de la información sobre el desempeño es la rendición de cuentas. Todos los países difunden los informes completos de las evaluaciones realizadas a través de internet, lo que permite a las organizaciones gubernamentales y no gubernamentales conocer la efectividad de la gestión de las intervenciones públicas. Algunos países publican no solamente los informes completos de las evaluaciones, sino también resúmenes que destacan los aspectos más importantes. No obstante, por lo general, esos

Existen muy pocos estudios sobre el uso efectivo de las evaluaciones y el grado en que estas sirven para mejorar la eficiencia técnica y asignativa de los recursos públicos.

resúmenes utilizan un formato y un lenguaje apropiado para especialistas y no para el público general. Conviene añadir que México (CONEVAL, 2017) y Perú (MEF, 2017) difunden, además, informes anuales sobre el seguimiento a los acuerdos de cambios a las intervenciones públicas.

Por otro lado, solamente Chile ha establecido procedimientos para que el Poder Legislativo use las evaluaciones como insumo para la discusión presupuestaria. La DIPRES envía los informes de las evaluaciones a la Comisión Permanente de Presupuestos del Congreso Nacional como parte del proceso que siguen los Protocolos de Acuerdo mencionados anteriormente. Además, en el Congreso de Chile existe la Unidad de Asesoría Presupuestaria que analiza asuntos presupuestarios y brinda asesoría a los congresistas en esa materia. Una de sus funciones es participar en la selección de los programas que la DIPRES evalúa cada año como parte de los *Protocolos de Acuerdo* y analizar los informes de esas evaluaciones. Tanto el Poder Ejecutivo como el Legislativo examinan los informes de las evaluaciones y los Compromisos Institucionales, durante todas las etapas del proceso de preparación y asignación presupuestaria, y toman decisiones que afectan el financiamiento de las intervenciones públicas evaluadas (Banco Mundial, 2005).

En el caso de México, aunque los informes de evaluación se envían a la Cámara de Diputados y el Congreso de la Unión de México también tiene una unidad de análisis presupuestario (Centro de Estudios de las Finanzas Públicas) que puede usar las evaluaciones para elaborar los análisis que le solicitan los congresistas, no existe un procedimiento específico para tal fin.

Existen muy pocos estudios sobre el uso efectivo de las evaluaciones y el grado en que estas sirven para mejorar la eficiencia técnica y asignativa de los recursos públicos. Chile realizó una evaluación de impacto del programa de evaluaciones de la DIPRES (Banco Mundial, 2005), Perú una evaluación de proceso de las EDEP que realiza el MEF (Baanante y Gracia, 2015), y México una evaluación externa del CONEVAL (Backhaus, Bamberger y Gómez, 2019).

En el primer caso, la evaluación indica que la DIPRES logró establecer “un programa complejo y eficaz y que, según la percepción de los participantes en el mismo, el número y la calidad de las evaluaciones y, sobre todo, sus usos prácticos, produce buenos resultados y se asegura que se apliquen” (Banco Mundial, 2005: 26). No obstante,

este y otros estudios posteriores señalan algunas limitaciones del sistema de evaluación chileno: centralización, excesivo énfasis en aspectos presupuestarios, y escaso diálogo con los organismos responsables de las intervenciones públicas evaluadas (Dussauge Laguna, 2011; Olavarría, 2012).

En relación con Perú, la evaluación señala que las evaluaciones han contribuido “en cierta medida, al logro de los objetivos de las intervenciones públicas a partir de las recomendaciones brindadas” y que los compromisos asumidos por las intervenciones públicas muchas veces “son inviables y los plazos considerados irreales” (Baanante y Gracia, 2015).

En cuanto a México, el informe de evaluación pone de relieve el alto rigor técnico, la alta credibilidad y la independencia del sistema encabezado por el CONEVAL, y señala que debe mejorar el uso y la influencia de la información que genera (Backhaus, Bamberger y Gómez, 2019). Además, en México se realizó un estudio académico sobre el uso de las evaluaciones en el sector de educación que concluyó que “La evaluación de los programas federales, como se encuentra actualmente diseñada, muestra una influencia limitada para cambios instrumentales con impacto en el desempeño. A pesar de que se han tenido avances importantes en la generación de evidencia con la instrumentación del sistema de evaluación de programas sociales, se vuelve necesario estudiar con mayor profundidad la puesta en marcha de la evaluación y, particularmente, el proceso de interacción y apropiación de la evidencia entre los actores relevantes” (Pérez Yarahuán, 2015: 708).

Hacen falta más estudios sobre el uso de las ERIP, especialmente en las siguientes áreas: i) el uso efectivo de las evaluaciones para mejorar la asignación de recursos públicos, ii) la aplicación efectiva de las recomendaciones de las evaluaciones, y iii) el efecto de los planes de mejora en la gestión de las intervenciones públicas. Los informes que emiten los países (CONEVAL, 2018; DIPRES, 2018; MEF, 2017) se basan en el seguimiento al cumplimiento de los acuerdos entre el ente rector de las evaluaciones y el organismo que ejecuta la intervención pública. Estos informes proporcionan datos cuantitativos sobre el porcentaje de cumplimiento de los acuerdos, pero no profundizan en las cualidades de las acciones ni en el efecto que han tenido en la eficiencia y eficacia de las intervenciones públicas, por lo tanto, no son suficientes para valorar los cambios que pueden atribuirse a los planes de mejora.

3. Recomendaciones para mejorar la gestión del desempeño mediante las ERIP

Crear sistemas nacionales de evaluación es una tarea compleja y de largo aliento. Varios autores han desarrollado manuales técnicos con los instrumentos y metodologías que pueden emplearse con ese fin (Gorgens y Kusek, 2010; Morra, Imas y Rist, 2009) y otros han analizado las características de estos sistemas tanto en América Latina (García Moreno y Feinstein, 2015; Pérez Yarahuán y Maldonado, 2015) como en otras regiones (Rist, Boily y Martin, 2011; Mackay, 2007). Esta sección pretende aportar algunos criterios que el autor considera importantes para fortalecer el uso de las ERIP en los países que las implementan y para iniciar su ejecución en los países que aún no lo hacen.

Recomendaciones para los países que ejecutan ERIP

Estandarización y flexibilización

La estandarización de la metodología de evaluación es propia de las ERIP y tiene varias ventajas: 1) permite incluir las evaluaciones de manera sistemática en el ciclo presupuestario que se caracteriza por rutinas altamente estructuradas e inflexibles; 2) abrevia la fase de diseño y preparación de las evaluaciones, aspecto crítico para que los informes de evaluación se produzcan oportunamente; 3) permite gestionar un mayor número de evaluaciones y, por lo tanto, obtener una mayor cobertura del gasto público sometido a evaluación; 4) facilita el seguimiento de la implementación de las evaluaciones y la revisión de los informes de evaluación; y 5) facilita el cálculo del costo de los estudios de evaluación. Estos aspectos son importantes para un buen funcionamiento del sistema de evaluación y han sido clave para el éxito de las EPG de la DIPRES, cuyo modelo se ha diseminado en ALC.

No obstante, la estandarización de la metodología se lleva a cabo a costa de restar a las evaluaciones un mejor ajuste a las características particulares de las intervenciones públicas, lo cual puede derivar en estudios sin la suficiente pertinencia. Algunos autores señalan que los responsables de las intervenciones públicas ven a la estandarización como un defecto de las ERIP (Baanante y Gracia, 2015; Banco Mundial, 2005; Backhaus, Bamberger y Gómez, 2019).

La flexibilización del modelo estandarizado podría realizarse adoptando tres medidas. En primer lugar, se sugiere elaborar una sección de preguntas de evaluación específica para cada intervención que

tome en cuenta las preocupaciones e intereses de sus gestores. En segundo lugar, se plantea eliminar las preguntas del modelo estandarizado que no sean relevantes para la evaluación o que no puedan responderse adecuadamente, debido a la falta de información u otras circunstancias. En tercer lugar, se requiere abrir un espacio para que los evaluadores formulen valoraciones sobre aspectos no previstos en el modelo estandarizado, pero que a su juicio son importantes para evaluar el desempeño, y sobre los que existe la información necesaria para estudiarlos.

Centralización y participación de los involucrados

Una de las características que diferencia a las ERIP de otros modelos de gestión de la evaluación es su manejo centralizado, lo que contribuye a que el ente encargado tenga autoridad y capacidad para gestionarlas. Su autoridad radica en que la ley le otorga la rectoría nacional del sistema, posición desde la cual puede injerir en todas las intervenciones públicas. Su capacidad consiste en los recursos financieros, técnicos y humanos que dispone para hacer funcionar el sistema. Ningún ministerio sectorial cumple con esas condiciones.

La centralización, sin embargo, carece de un elemento importante para promover la gestión del desempeño: la aceptación de los involucrados. En efecto, es común que los involucrados en la ejecución de las intervenciones públicas acepten con reservas o rechacen abiertamente los procedimientos e instrumentos de las ERIP, según se desprende de los estudios que han observado este aspecto (Baanante y Gracia, 2015; Olavarría, 2012). En muchos casos la aceptación se reduce al acatamiento de las normas, pero no se produce una contribución activa en el proceso. En circunstancias como estas, la probabilidad de usar los resultados de la evaluación como medio para mejorar el desempeño es baja.

Un mayor involucramiento de los organismos ejecutores de las intervenciones públicas en la selección de las intervenciones que se evaluarán anualmente, en la elaboración de los términos de referencia de la evaluación, en la selección de los evaluadores, en la revisión de los informes de evaluación y en la redacción de los acuerdos de mejora que se derivan de dicho informe, podría contribuir a una mayor aceptación del ejercicio centralizado de las ERIP.

Una opción más radical consiste en adoptar el camino de México y crear un ente de evaluación independiente, que no tenga el sesgo de los entes rectores del presupuesto o la planificación, y que opere

con un mayor grado de coordinación con todos los involucrados. No obstante, esta opción -que afecta el diseño de todo el sistema de evaluación y que rebasa ampliamente el campo de las ERIP- es apta para países con un grado relativamente alto de capacidad institucional en evaluación como Chile o Colombia¹⁵.

Gestión del diálogo técnico

Una buena evaluación tiene potencial para promover un diálogo sustancioso entre los actores involucrados en el financiamiento, la ejecución y el control de las intervenciones públicas. Ese diálogo empieza con la selección de lo que se va a evaluar y llega a su punto culminante con las discusiones acerca de los resultados del estudio y de las medidas que se tomarán para superar los problemas detectados: el diálogo cumple la función de precursor del uso de las evaluaciones.

La conducción del diálogo técnico es una de las responsabilidades de la entidad que gestiona el sistema y debe llevarse a cabo con la rigurosidad de los otros procesos del ciclo de evaluación. Por lo general, este es un aspecto poco considerado en las normas, procedimientos y prácticas que rigen las evaluaciones. Si bien los países estudiados han establecido los productos que resultan de ese diálogo (plan anual de evaluaciones, plan de mejoras de la intervención pública evaluada), conviene ampliar y profundizar el tratamiento y los mecanismos del diálogo técnico y concebirlos en el marco de una estrategia de gestión del conocimiento.

Moynihan (2008) propone que los diálogos sobre el desempeño sean periódicos y formen parte de una rutina previamente establecida; que se conduzcan con mecanismos de facilitación y reglas claras, mediante un enfoque cooperativo; que se basen tanto en la información sobre el desempeño como en el conocimiento y la experiencia sobre los éxitos, fracasos e innovación en un ámbito dado; y que se examine no solo el cumplimiento de objetivos, sino también los supuestos que los sostienen.

Uso de las evaluaciones para optimizar la asignación presupuestaria

El uso de las evaluaciones para mejorar la asignación presupuestaria es uno de los elementos más débiles del presupuesto por resultados, incluso en los países pioneros en su aplicación como Australia y Estados Unidos (Robinson, 2014; Moynihan, 2008). En parte, esto

Una de las limitaciones que frecuentemente encuentran los evaluadores en ALC es la calidad deficiente y/o la ausencia de los registros administrativos necesarios para valorar la implementación y los resultados de las intervenciones públicas.

se debe a la creencia extendida de que el sistema de indicadores es suficiente para analizar el desempeño.

Robinson (2014) sugiere que para aprovechar mejor las evaluaciones es necesario dirigir las hacia los programas que pueden recortarse o eliminarse y enfocarse en los elementos que más relevancia tienen para recortar los gastos. Plantea que conviene partir de evaluaciones de desempeño y/o diseño que informen sobre la pertinencia de continuar o no financiando las intervenciones públicas. Sugiere que se reserven las evaluaciones de procesos para las intervenciones que muestran un mal desempeño, pero cuya evaluación de diseño es favorable. En este contexto, las evaluaciones de procesos servirían para determinar si los recursos que se requieren para mejorar la intervención pública son aceptables desde un punto de vista económico.

Fortalecer el sistema de seguimiento

Una de las limitaciones que frecuentemente encuentran los evaluadores en ALC es la calidad deficiente y/o la ausencia de los registros administrativos necesarios para valorar la implementación y los resultados de las intervenciones públicas. Este problema afecta la validez de los indicadores que alimentan los sistemas de seguimiento, uno de los principales insumos de las ERIP.

La solidez de los sistemas de información y de los sistemas de indicadores del desempeño que de ellos se derivan, es clave para la producción de evaluaciones de calidad. El desarrollo y perfeccionamiento de esos sistemas es un desafío de largo aliento que compete a múltiples actores. No obstante, el órgano encargado del sistema de evaluación puede impulsar auditorías de datos que detecten problemas y propongan soluciones que contribuyan a mejorar la validez de los indicadores de desempeño. Esta es una acción que puede realizarse en cooperación con la oficina nacional de estadísticas.

Analizar el efecto de los acuerdos de mejora a las intervenciones públicas

Aunque cabría esperar esfuerzos evaluativos de su propia acción por parte de organizaciones encargadas de fomentar y gestionar evaluaciones, esta no es una práctica común. En particular, hacen falta estudios sobre los efectos que tienen los acuerdos de mejoramiento que surgen de las evaluaciones. Solamente dos acciones se han realizado en este campo. Una fue la evaluación ya citada de

las EPG que la DIPRES encargó al Banco Mundial en 2005. La otra consistió en una iniciativa que la Secretaría de la Función Pública de México llevó a cabo, en colaboración con el Centro de Investigación y Docencia Económicas (CIDE) y la asociación civil México Evalúa, para fomentar el uso de las evaluaciones (Maldonado y Cejudo, 2011). Estas organizaciones analizaron la manera en que varios programas federales implementaron las acciones de mejora derivadas de los resultados de las evaluaciones.

Se sugiere, por lo tanto, emprender estudios que permitan determinar el grado en que los planes de mejoramiento efectivamente influyen en la gestión de las intervenciones públicas, trascendiendo el enfoque de seguimiento que hasta ahora ha primado en todos los países. Los estudios deberían realizarse en el marco de una estrategia bien concebida de diálogo técnico e incluso ser acompañados con un mecanismo de incentivo como el ensayo en México.

Recomendaciones para los países que no ejecutan ERIP

La mayoría de los países de ALC no tiene un sistema nacional de evaluación de desempeño de las intervenciones públicas, aunque cuentan con leyes generales que establecen la evaluación como una de las fases del ciclo de gestión de las políticas públicas (García Moreno y Feinstein, 2015). La primera recomendación para esos países -y la más obvia- es emprender la institucionalización de las ERIP como parte de un esfuerzo integral orientado a mejorar la gestión del desempeño. La responsabilidad de la gestión de las evaluaciones puede recaer tanto en el ente rector del presupuesto como en el de la planificación, o en una unidad encargada de impulsar el logro de las prioridades fijadas por el gobierno¹⁶.

Evaluar en el marco de una estrategia de gestión del desempeño

El desarrollo de las ERIP debe inscribirse en una estrategia para implementar diversos mecanismos de gestión del desempeño. Las actividades de evaluación que no están ancladas en dicha estrategia derivan, tarde o temprano, en meros trámites administrativos que no inciden positivamente en el quehacer de las intervenciones públicas y que, con frecuencia, se ganan mala reputación entre las autoridades y los funcionarios que los ven como una carga de trabajo inútil. Una estrategia de esa índole supone desarrollar sistemas y mecanismos que actúen de manera coordinada en los cinco pilares de la gestión para resultados: planificación, presupuesto, gestión

Las normas, los procedimientos y los instrumentos que dan vida al sistema de evaluación deben construirse teniendo en mente el objetivo principal de mejorar el desempeño de la gestión pública y el uso que se dará a los estudios evaluativos, es decir, rendición de cuentas, control y dirección, y aprendizaje organizacional.

financiera pública, gestión de programas y proyectos, seguimiento y evaluación (García López y García Moreno, 2010).

Este enfoque supone que las normas, los procedimientos y los instrumentos que dan vida al sistema de evaluación deben construirse teniendo en mente el objetivo principal de mejorar el desempeño de la gestión pública y el uso que se dará a los estudios evaluativos, es decir, rendición de cuentas, control y dirección, y aprendizaje organizacional (Van Dooren, Halligan y Bouckaert, 2010). Aunque estos usos se complementan mutuamente, en la práctica podría no ser sencillo armonizarlos, porque los actores que participan en la gestión pública tienen intereses diversos y lo que interesa a unos no necesariamente interesa a otros.

Así, por ejemplo, el ente rector del presupuesto podría estar interesado en usar las evaluaciones para restringir los gastos, mientras las organizaciones que ejecutan las intervenciones públicas podrían querer emplearlas para justificar un incremento de financiamiento. Del mismo modo, el congreso y las entidades superiores de fiscalización priorizan la información que les permite cumplir su función de control, en tanto las autoridades ministeriales se interesan en mejorar la capacidad de gestión de las entidades para lograr los objetivos gubernamentales.

Por esta razón, conviene que el diseño inicial del sistema considere los puntos de vista y necesidades de información de todos los actores y no solamente los requerimientos del ente que administra las evaluaciones, pues la sostenibilidad y afianzamiento del sistema depende de la manera en que se logre armonizar los intereses de los actores y combinar los usos posibles.

Aprender y no adoptar

Aunque la experiencia internacional ayuda a bosquejar la estructura de un sistema de evaluación y es útil como referente, son las características propias de un país -sistema político, marco legal e institucional, capacidades organizacionales- las que determinan el contenido y la ruta del proceso de construcción. La adopción de modelos que han sido exitosos en otros países es un camino que frecuentemente lleva al fracaso en los intentos por mejorar la gestión pública (Andrews, Pritchett y Woolcock, 2017). Aprender de otras experiencias y adaptar críticamente modelos e instrumentos de otros a las necesidades y características propias es más fructífero que adoptarlas¹⁷. Los sistemas de Chile y México son ejemplos de este enfoque de construcción.

La idea de que no hay un modelo a seguir, sino más bien propósitos que cumplir y problemas que resolver, permite emprender la tarea con plena conciencia de su complejidad y con un espíritu creativo e innovador. Paralelamente, esto requiere una actitud de vigilancia constante de lo que funciona y no funciona, de ensayo y error, de aprendizaje continuo.

Fortalecer la capacidad institucional

La tarea más compleja, costosa y de largo aliento que enfrentan quienes ponen en marcha un sistema nacional de evaluación es la creación de las capacidades institucionales necesarias que giran alrededor de tres funciones clave del sistema: 1) gestión del sistema de evaluación, 2) elaboración de los estudios de evaluación y 3) uso de las evaluaciones.

La gestión del sistema supone adquirir la capacidad para administrar el proceso de producción de las evaluaciones, difundir apropiadamente los estudios a distintos públicos, facilitar el diálogo sobre los resultados de las evaluaciones y hacer seguimiento al cumplimiento de los acuerdos que surgen de ese diálogo. El órgano que está a cargo de la gestión del sistema necesita crear los procesos e instrumentos correspondientes y capacitar a sus funcionarios para que se encarguen de ellos.

En segundo lugar, se requieren evaluadores competentes y experimentados. En los países en los que el Estado no evalúa sistemáticamente las intervenciones públicas, el mercado de consultoría de evaluaciones es incipiente y las organizaciones públicas no tienen a su alcance suficientes profesionales con experiencia en este campo. La asistencia técnica internacional y la contratación de evaluadores extranjeros es una solución temporal que puede funcionar al inicio de la implementación del sistema, pero que no es sustentable a mediano y largo plazo. Por eso, es necesario diseñar una estrategia que permita generar un cuadro de profesionales nacionales capaces de operar el sistema, sea como funcionarios públicos o como consultores independientes. Conviene que las universidades nacionales o internacionales construyan una oferta académica permanente en este campo, integrando profesores y desarrollando cursos de corto, mediano y largo alcance. También conviene apoyar la constitución de redes de expertos en evaluación para que compartan sus técnicas, metodologías y experiencias.

En tercer lugar, el uso efectivo de las evaluaciones demanda que todos los funcionarios públicos involucrados aprovechen los

En ALC la adopción generalizada de sistemas de indicadores para hacer seguimiento a la ejecución del presupuesto y/o a los planes de gobierno, por lo general, no se ha complementado con iniciativas vigorosas que fomenten la evaluación de las intervenciones públicas, lo cual limita seriamente la utilidad de tales sistemas.

estudios evaluativos para los diversos propósitos y usos que tiene la evaluación. Esta capacidad no se obtiene solo mediante actividades específicas de formación o capacitación individual, sino que es el fruto de un ambiente habilitador compuesto por las normas, los procedimientos, los conocimientos y las prácticas que componen la cultura de las organizaciones del sector público y que van formándose con el pasar de los años. Por esta razón, es necesario impulsar la implementación sostenida de reformas orientadas a mejorar el desempeño de la gestión pública.

Puesto que la capacidad para usar efectivamente los resultados de las evaluaciones rebasa ampliamente el ámbito de incidencia del ente que administra el sistema, es conveniente que este promueva mecanismos específicos -como los planes de mejoramiento- para asegurar que eso ocurra. Al mismo tiempo, debe generar reglas que eviten el incumplimiento de los acuerdos e incentiven la buena gestión del desempeño.

Conclusiones

En ALC la adopción generalizada de sistemas de indicadores para hacer seguimiento a la ejecución del presupuesto y/o a los planes de gobierno, por lo general, no se ha complementado con iniciativas vigorosas que fomenten la evaluación de las intervenciones públicas, lo cual limita seriamente la utilidad de tales sistemas. Sin embargo, existe un pequeño grupo de países que realizan sistemáticamente evaluaciones, entre los que se encuentran Chile, Colombia, México, Paraguay, Perú y Uruguay. Para ello han creado sistemas de evaluación ligados a las funciones de planificación y/o presupuesto nacional que generalmente realizan evaluaciones rápidas y evaluaciones de impacto.

Las ERIP que se llevan a cabo en ALC tienen dos propósitos: mejorar el diseño, calidad, eficiencia y eficacia de las intervenciones públicas, e informar a los actores que intervienen en el ciclo presupuestario acerca del desempeño de las intervenciones públicas. Las características más importantes de este tipo de evaluación son cuatro: 1) se gestionan de manera centralizada, 2) usan una metodología estandarizada, 3) se ejecutan en lapsos cortos a costos bajos y 4) utilizan la información sobre el desempeño que ya tienen los programas.

El ciclo de ejecución de las ERIP tiene tres momentos: la selección de las intervenciones públicas que se evaluarán, la preparación y ejecución de los estudios de evaluación, y el uso y difusión de los

resultados de las evaluaciones. En los países que implementan un presupuesto por resultados ese ciclo está articulado con las etapas del ciclo presupuestario: programación, formulación, discusión y aprobación, ejecución y evaluación. El proceso que más desafíos presenta en la mayoría de los países es el uso de las evaluaciones. No obstante, salvo casos puntuales, no se realizan estudios sobre el uso efectivo de las evaluaciones para mejorar la gestión de las intervenciones públicas y para optimizar la asignación de los recursos.

Además, se requiere fomentar estudios acerca de la efectividad de los planes de mejoramiento que los entes rectores del sistema de evaluación acuerdan con los ejecutores de las intervenciones públicas, de la calidad del diálogo técnico entre los actores que participan en las evaluaciones, y de la efectividad de los mecanismos ligados a las ERIP que algunos países usan para mejorar la gestión del desempeño.

Las recomendaciones para los países que ya realizan ERIP son: 1) incluir elementos que flexibilicen la metodología estandarizada de las ERIP para dar cabida a las preguntas de los ejecutores de las intervenciones que se evalúan y para que se adapte mejor a las características del objeto evaluado; 2) propiciar una mayor participación de los actores institucionales involucrados en la evaluación; 3) crear procedimientos para la gestión del diálogo técnico en torno a la selección, implementación y uso de las evaluaciones; 4) desarrollar criterios para usar las evaluaciones como medio para mejorar la asignación presupuestaria; 5) fortalecer la calidad de los datos de los sistemas de información y promover la mejora continua de los sistemas de indicadores de desempeño; y 6) analizar el efecto de los acuerdos de mejora a las intervenciones públicas.

Las recomendaciones para los países que aún no implementan ERIP son: 1) introducir las ERIP en el marco de una estrategia de gestión del desempeño; 2) analizar los modelos que otros países han seguido en la construcción de sus sistemas de evaluación y adaptarlos a las necesidades y características propias, en lugar de adoptarlos automáticamente; y 3) fortalecer la capacidad institucional para gestionar, implementar y usar las evaluaciones.

Notas

(1) En este texto se usa el término *intervenciones públicas* para abarcar las políticas, programas, proyectos, medidas administrativas y otras actividades gubernamentales tendientes a generar valor público.

(2) Se usará el término *presupuesto por resultados* en lugar de *presupuesto basado en el desempeño* debido a que, aunque este último es más preciso, el primero tiene una mayor difusión en ALC.

(3) El cuadro presenta la última actualización de las normas. En varios casos, las normas originales son más antiguas y también menos detalladas.

(4) Tomado de <http://www.mef.gob.pe/es/presupuesto-por-resultados/instrumentos/evaluaciones-independientes/655-presupuesto-publico/5356-evaluaciones-de-diseno-y-ejecucion-presupuestal-edep> (visitado el 11 de enero de 2019).

(5) El DNP de Colombia realizó evaluaciones Ejecutivas entre 2006 y 2011.

(6) El Ministerio de Hacienda de Paraguay realizó evaluaciones de Diseño y Gestión entre 2011 y 2017.

(7) El Ministerio de Hacienda de Paraguay realizó evaluaciones de Eficiencia entre 2011 y 2017.

(8) El autor de este artículo tuvo acceso solamente a los datos de Chile respecto a la relación entre los programas e instituciones evaluados y el presupuesto. No se encontró información pública sobre este aspecto en los otros países.

(9) Datos extraídos del *Repositorio de evaluaciones* de SINERGIA el 19-09-2018: <http://sinergiapp.dnp.gov.co/#Evaluaciones/EvalFin>.

(10) La evaluación de los programas a cargo de otros sectores es función de la Secretaría de Hacienda y Crédito Público.

(11) Información extraída de la página web de Transparencia Presupuestaria de la SHCP el 02-11-2020: <https://www.transparenciapresupuestaria.gob.mx/es/PTP/evaluaciones>.

(12) Información extraída de la página de Internet del MEF el 20-09-2019: <https://www.mef.gob.pe/es/presupuesto-por-resultados/instrumentos/evaluaciones-independientes>.

(13) Datos extraídos del repositorio de evaluaciones del Portal de Transparencia Presupuestaria el 28-03-2019: <https://transparenciapresupuestaria.opp.gub.uy/inicio/registro-nacional-de-evaluaciones/busqueda>.

(14) Las razones que explican esta conducta son de diversa índole: el carácter intrínsecamente político de la asignación de los recursos públicos, las dificultades de articular la lógica multianual de las políticas públicas con la lógica anual del presupuesto, la incipiente implementación de la gestión para resultados en las organizaciones públicas, la debilidad de los sistemas de Gestión Financiera Pública, la limitada cantidad de productos (bienes y servicios) proclives a financiarse mediante mecanismos que ligan la asignación de los recursos con los resultados (financiamiento por fórmula, incentivos por cumplimiento de metas comprador-proveedor), la gestión inadecuada del proceso de reformas, etc.

(15) Esta opción se ha discutido en Chile desde hace algunos años (Bellolio, Morales y Sierpe, 2012; Olavarría Gambi y Figueroa Huencho, 2010).

(16) A estas unidades también se las conoce con el nombre de *Centro de Gobierno*. En inglés se las llama *Delivery Unit*.

(17) Se usa el término adoptar en la siguiente acepción del Diccionario de la Lengua Española: recibir, haciéndolo propio, un parecer, un método, una doctrina, etc., que han sido creados por otros.

Bibliografía

- Allen, Richard y Tomassi, Daniel (eds.) (2001), *Managing Public Expenditure. A Reference Book for Transition Countries*, Paris, OECD.
- Andrews, Matt; Pritchett, Lant; y Woolcock, Michael (2017), *Building State Capability: Evidence, Analysis, Action*, Oxford, Oxford University Press.
- Anker, M.; Guidotti, R. J.; Orzeszyna, S.; Sapirie, S. A.; y Thuriaux, M. C. (1993), "Rapid Evaluation Methods (REM) of Health Services Performance: Methodological Observations", en *Bulletin of the World Health Organization*, Vol. 71 N° 1, pp. 15-21.
- Arancibia, Emilia (2018), "Evaluación de programas e instituciones públicas de la dirección de presupuestos en el Ministerio de Educación", Santiago, Ministerio de Educación.
- Arana Araya, Ignacio (2013), "Informal Institutions and Horizontal Accountability: Protocols in the Chilean Budgetary Process", en *Latin American Politics and Society*, Vol. 55 N° 4, pp. 74-94.
- Baanante, Miguel Jaramillo y Gracia, Orlando (2015), "Evaluación del proceso de evaluación de las intervenciones públicas a través del instrumento EDEP en el marco del Presupuesto por Resultados", Lima, Ministerio de Economía y Finanzas; Centro Regional para el Aprendizaje en Evaluación y Resultados para América Latina y el Caribe; Secretaría de Estado para Asuntos Económicos de Suiza; Agencia Alemana de Cooperación Internacional.
- Backhaus, Annette; Bamberger, Michael; y Gómez, Ricardo (2019), "Evaluación externa: Consejo Nacional de Evaluación de la Política de Desarrollo Social (CONEVAL), México", México, Agencia Alemana de Cooperación Internacional, CONEVAL.
- Banco Mundial (2005), "Chile: estudio de evaluación de impacto del Programa de Evaluación de Programas; resumen ejecutivo", Washington, Banco Mundial. Unidad de Reducción de la Pobreza y Gestión Económica en América Latina y el Caribe (Informe N° 34589-CL. Volumen I: Informe principal).
- Belloio, Álvaro; De Los Ríos, Boris; Irrázaval, Ignacio; Larraín, Luis; Marshall, Jorge; Morales, Juan José; Sierpe, Emilio; y Véliz, Valeska (2012), "Agencia de Evaluación Pública: propuesta de diseño institucional", Santiago, s.e.

- CEPAL (2018), *Panorama de la gestión pública en América Latina y el Caribe. Un gobierno abierto centrado en el ciudadano*, Santiago, Naciones Unidas; CEPAL; ILPES.
- CONEVAL (2011), "Mecanismo para el seguimiento a los aspectos susceptibles de mejora derivados de informes y evaluaciones a los programas presupuestarios de la Administración Pública Federal", México, Consejo Nacional de Evaluación de la Política de Desarrollo Social.
- _____ (2015), "La política de evaluación en México: 10 años del CONEVAL", México, Consejo Nacional de Evaluación de la Política de Desarrollo Social.
- _____ (2017) "Informe de seguimiento a los aspectos susceptibles de mejora de programas y acciones federales de desarrollo social 2016-2017: balance anual sobre el uso de las evaluaciones", México, Consejo Nacional de Evaluación de la Política de Desarrollo Social.
- _____ (2018), "Informe de seguimiento a los aspectos susceptibles de mejora de los programas y acciones federales de desarrollo social 2017-2018: balance anual sobre el uso de las evaluaciones", México, Consejo Nacional de Evaluación de la Política de Desarrollo Social.
- CONEVAL, SHCP Y SFP (2007), "Lineamientos generales para la evaluación de los programas federales de la Administración Pública Federal", México, Consejo Nacional de Evaluación de la Política de Desarrollo Social; Secretaría de Hacienda y Crédito Público; Secretaría de la Función Pública.
- CONPES y DNP (2004), "Renovación de la administración pública: gestión por resultados y reforma del Sistema Nacional de Evaluación", Bogotá, Consejo Nacional de Política Económica y Social; Departamento Nacional de Planeación (Documento CONPES; N° 3294).
- Darville, Paula; Díaz, Rodrigo; y Leiva, Juan P. (2017), "Cobertura evaluación de programas e instituciones públicas", Santiago, Ministerio de Hacienda. Dirección de Presupuestos.
- DIPRES (s.f), "Pauta para elaboración de informe final de evaluación", Santiago, Ministerio de Hacienda. Dirección de Presupuestos. División de Control de Gestión.
- _____ (2007), "Notas técnicas", Santiago, Ministerio de Hacienda. Dirección de Presupuestos.
- _____ (2015), "Evaluación ex-post: conceptos y metodologías", Santiago, Ministerio de Hacienda. Dirección de Presupuestos. División de Control de Gestión.
- _____ (2016), "Evaluación focalizada de ámbito EFA", Santiago, Ministerio de Hacienda. Dirección de Presupuestos.

- _____ (2018), "Informe de seguimiento del cumplimiento de compromisos 2018", Santiago, Ministerio de Hacienda. Dirección de Presupuestos.
- DNP (2010), "15 años del Sistema de Seguimiento y Evaluación en Colombia - Sinergia: una mirada desde las evaluaciones de política pública más relevantes", Bogotá, Departamento Nacional de Planeación (Serie 1: Evolución de Sinergia y Evaluaciones en Administración del Estado).
- Dussauge Laguna, Mauricio (2011), "Chile's M&E System, 1994-2010", en *Building Better Policies. The Nuts and Bolts of Monitoring and Evaluation Systems*, Gladys López, Phillip Krause and Keith Mackay (eds.), Washington, World Bank.
- Feinstein, Osvaldo (2016), "Rapid Evaluations", Washington, World Bank.
- García López, Roberto y García Moreno, Mauricio (2010), *La gestión para resultados en el desarrollo: avances y desafíos en América Latina y el Caribe*, Washington, BID.
- García Moreno, Mauricio y Feinstein, Osvaldo (2015), "Seguimiento y evaluación", en *Construyendo gobiernos efectivos: logros y retos de la gestión pública para resultados en América Latina y el Caribe*, Jorge Kaufmann, Mario Sanginés y Mauricio García Moreno (eds.), Washington, BID.
- García Moreno, Mauricio; Kaufmann, Jorge; y Sanginés, Mario (eds.) (2015), *Construyendo gobiernos efectivos: logros y retos de la gestión pública para resultados en América Latina y el Caribe*, Washington, BID.
- Gorgens, Marelize y Zall Kusek, Jody (2010), *Making Monitoring and Evaluation Systems Work: a Capacity Development Tool Kit*, Washington, World Bank.
- Hargreaves, Margaret B. (2014), "Rapid Evaluation Approaches for Complex Initiatives; White Paper", Cambridge, Mathematica Policy Research.
- Mackay, Keith (2007), *Cómo crear sistemas de SyE que contribuyan a un buen gobierno*, Washington, Banco Mundial.
- Maldonado, Claudia y Guillermo Cejudo (eds.) (2011), *De las recomendaciones a las acciones: la experiencia del Premio 2011 Programas Federales Comprometidos con el Proceso de Evaluación*, México, Centro Regional para el Aprendizaje en Evaluación y Resultados para América Latina y el Caribe; Centro de Investigación y Docencia Económicas.
- McNall, Miles y Foster-Fishman, Pennie G. (2007), "Methods of Rapid Evaluation, Assessment, and Appraisal", en *American Journal of Evaluation*, Vol. 28 N° 2, pp. 151-168.

- MEF (2008), "Directiva para las evaluaciones independientes en el sistema nacional de presupuesto público en el marco del presupuesto por resultados (Directiva No 009-2008-EF/76.01 modificada por la Resolución Directoral No 023-2012-EF/50.01)", Lima, Ministerio de Economía y Finanzas.
- _____ (2017), "Informe de seguimiento del cumplimiento de compromisos de las Evaluaciones de Diseño y Ejecución Presupuestal (EDEP) al 31 de diciembre de 2017", Lima, Ministerio de Economía y Finanzas. Dirección General de Presupuesto Público.
- _____ (2018), "Gestión de finanzas públicas y presupuesto por resultados. Concept Note sobre asistencia técnica requerida para implementar las recomendaciones planteadas por la OECD Public Governance Review", Lima, Ministerio de Economía y Finanzas.
- MIDEPLAN (2012), "Manual gerencial para el diseño y ejecución de evaluaciones estratégicas de gobierno", San José, Ministerio de Planificación Nacional y Política Económica. Área de Evaluación y Seguimiento.
- _____ (2017), "Manual de evaluación para intervenciones públicas", San José, Ministerio de Planificación Nacional y Política Económica. Área de Evaluación y Seguimiento.
- Ministerio de Hacienda (2003), "Decreto 1.177 del 17 de diciembre. Reglamento para el funcionamiento del sistema de evaluación de programas establecido en el Artículo 52 del Decreto Ley No 1.263 de 1975, Orgánico de Administración Financiera del Estado", Santiago, Ministerio de Hacienda. Dirección de Presupuestos.
- Morra Imas, Linda G. y Rist, Ray (2009), *The Road to Results: Designing and Conducting Effective Development Evaluations*, Washington, World Bank.
- Moynihan, Donald (2008), *The Dynamics of Performance Management: Constructing Information and Reform*, Washington, Georgetown University Press. 12. ed.
- _____ (2013), "Advancing the Empirical Study of Performance Management: What We Learned from the Program Assessment Rating Tool", en *American Review of Public Administration*, Vol. 43 N° 5, pp. 499-517.
- Moynihan, Donald y Beazley, Ivor (2016), *Toward Next-Generation Performance Budgeting: Lessons from the Experiences of Seven Reforming Countries*, Washington, World Bank.
- Olavarría Gambi, Mauricio (2012), "La evaluación de programas en Chile: análisis de una muestra de programas públicos evaluados", en *Revista del CLAD Reforma y Democracia*, N° 54, octubre, pp. 143-178.

- Olavarría Gambi, Mauricio y Figueroa Huencho, Verónica (2010), *Una agencia de evaluación de políticas públicas para Chile: lecciones de la historia y de la experiencia internacional*, Santiago, Editorial Universitaria.
- Pérez Yarahuán, Gabriela (2015), "La influencia de la evaluación externa en los programas gubernamentales para la educación básica, 2002-2012", en *Revista Mexicana de Investigación Educativa*, Vol. 20 N° 66, pp. 685-710.
- Pérez Yarahuán, Gabriela y Maldonado, Claudia (eds.) (2015), *Panorama de los sistemas nacionales de monitoreo y evaluación en América Latina*, México, Centro Regional para el Aprendizaje en Evaluación y Resultados para América Latina y el Caribe; Centro de Investigaciones y Docencia Económicas.
- Rist, Ray C.; Boily, Marie-Helene; y Martin, Frédéric (eds.) (2011), *Influencing Change: Building Evaluation Capacity to Strengthen Governance*, Washington, World Bank.
- Robinson, Marc (2014), "Connecting Evaluation and Budgeting", Washington, World Bank. Independent Evaluation Group (ECD Working Paper Series; N° 30).
- Schick, Allen (2014), "The Metamorphoses of Performance Budgeting", en *OECD Journal on Budgeting*, Vol. 2013 N° 2, pp. 49-79.
- Van Dooren, Wouter; Bouckaert, Geert; y Halligan, John (2010), *Performance Management in the Public Sector*, Abingdon, Taylor and Francis.
- Weiss, Carol H. (1998), *Evaluation: Methods for Studying Programs and Policies*, Upper Saddle River, Prentice Hall. 2. ed.